



derived from ORESTES PCR (U.S. Letters Patent application No. 19,716 - Ludwig Institute for Cancer Research Profiles into the PUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match Score 34.6%; Best Local Similarity 37.0%; Matches 30; Conservative 22; Mismatches 29; Indels 0; Gaps 0;  
 Qy 12 AAUCAUCCAGGCAUAUGUCGACUCCGUUAAACCAAUAGGUUAAG 71  
 Db 150 AATAATTAGGGCTGTCTAGTACTTTGGTTCACAAATAAAATGGTTAAC 209  
 Qy 72 CAUAGGCCUTAGGCCACAGCAA 92  
 Db 210 AATTCGGTTACACACACAA 230

RESULT 2  
 LOCUS AG557770 DNA, clone:MSMg01-476B14.T7, genomic survey sequence.  
 DEFINITION Mus musculus molossinus DNA, clone:MSMg01-476B14.T7, genomic survey sequence.

ACCESSION AG557770  
 VERSION 0.1  
 KEYWORDS GSS.  
 SOURCE Mus musculus molossinus  
 ORGANISM Mus musculus molossinus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
 Unpublished (2000)  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: dduun@gene.tics.utah.edu

COMMENT Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-2-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.jp, URL:http://hp.gsc.riken.go.jp/ , Tel:81-45-503-9111, Fax:81-45-503-9170).  
 Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@tc.riken.jp).  
 Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1  
 Royada, Tsukuba, 305-0074 Japan.  
 Phone: 81-3-36-9189, fax: 81-298-36-9199  
 e-mail: abe@tc.riken.jp

BRIMERS Sequencing : T7

LIBRARY Vector : pBACE3.6

R.Site 1 : EcoRI

R.Site 2 : Location/Qualifiers

FEATURES source 1. /organism="Mus musculus molossinus"  
 /mol\_type="genomic DNA"  
 /sub\_species="molossinus"  
 /db\_xref="taxon:57486"  
 /clone="MSMg01-476B14.T7"  
 /sex="male"  
 /tissue type="mixture of kidney and spleen"  
 /clone\_Lib=MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match Score 31.2%; Best Local Similarity 32.6%; Matches 30; Conservative 24; Mismatches 38; Indels 0; Gaps 0;  
 Query Match Score 31.2%; Best Local Similarity 32.6%; Prod. No. 9,4; Matches 25; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

Query Match Score 41. UCCCCUCAAAACCAGUTAATAAUCCGUTUAGCAUAUGCCUUCGCGAAG 89  
 Best Local Similarity 51.0%; Prod. No. 8,6; Matches 25; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

Qy 41 UCCCCUCAAAACCAGUTAATAAUCCGUTUAGCAUAUGCCUUCGCGAAG 89  
 Db 539 TTTCCCTAAAACCACTTATAAAATCTTGCTAGCATATGCAATAATGTTCTG 491

RESULT 3

ACCESSION AZ955882  
 VERSION 1  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 442)  
 AUTHORS Dunn,D., Royagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meinen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.

TITLE Plasmid whole genome scaffolding with paired end reads from 10kb plasmid inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 Tel: 801 585 5606  
 Fax: 801 585 7177

ACCESSION AG557770  
 VERSION 0.1  
 KEYWORDS GSS.  
 SOURCE Mus musculus molossinus  
 ORGANISM Mus musculus molossinus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
 AUTHORS Title BAC and Sequences of Library MSMg01  
 JOURNAL Unpublished (2000)  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: dduun@gene.tics.utah.edu

ACCESSION AG557770  
 VERSION 0.1  
 KEYWORDS GSS.  
 SOURCE Mus musculus molossinus  
 ORGANISM Mus musculus molossinus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
 AUTHORS Title Direct Submission  
 JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-2-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.jp, URL:http://hp.gsc.riken.go.jp/ , Tel:81-45-503-9111, Fax:81-45-503-9170).

COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@tc.riken.jp).  
 Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1  
 Royada, Tsukuba, 305-0074 Japan.  
 Phone: 81-3-36-9189, fax: 81-298-36-9199  
 e-mail: abe@tc.riken.jp

BRIMERS Sequencing : T7

LIBRARY Vector : pBACE3.6

R.Site 1 : EcoRI

R.Site 2 : Location/Qualifiers

FEATURES source 1. /organism="Mus musculus molossinus"  
 /mol\_type="genomic DNA"  
 /sub\_species="molossinus"  
 /db\_xref="taxon:57486"  
 /clone="MSMg01-476B14.T7"  
 /sex="male"  
 /tissue type="mixture of kidney and spleen"  
 /clone\_Lib=MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match Score 31.2%; Best Local Similarity 32.6%; Prod. No. 9,4; Matches 30; Conservative 24; Mismatches 38; Indels 0; Gaps 0;

Query Match Score 31.2%; Best Local Similarity 32.6%; Prod. No. 9,4; Matches 30; Conservative 24; Mismatches 38; Indels 0; Gaps 0;



/clone="IMAGE:3586661"	Qy	22 GGCAGAUAUGHGACTCCGCUCCUCUCAACCAGUAAAATGGGTTTACCAUAUGCCUU	81
/tissue_type="spontaneous tumor, metastatic to mammary.	Db	366 GGTATCTGTCGTCCTCCTACGTGGGTATAATGGATTAAATACCTT	307
stem cell origin."			
/lab_host="PH10B"			
/clone_lib="NCI_CGAP_Lu29"			
/note="Organ: Lung; Vector: PCMV-SHORT6; Site:1: SaltI;	Qy	82 AGCG 85	
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.	Db	306 CCCG 303	
Library constructed by Life Technologies. Investigator			
providing samples: Gilbert Smith, NIH"			
ORIGIN			
Query Match Similarity 30.6%; Score 30.6; DB 2; Length 1074;			
Best Local Similarity 44.2%; Pred. No.17; Indels 0; Gaps 0;			
Matches 34; Conservative 14; Mismatches 29; Locus CK016754/c			
DEFINITION linear EST 26-NOV-2003			
JOURNAL IMAGE:7043617			
COMMENT mRNA clone IMAGE:NIH_ZGC_10 Danio rerio cDNA clone IMAGE:NIH_ZGC_10			
ORGANISM Danio rerio			
KEYWORDS zebrafish			
REFERENCE CK016754.1 GI:38541867			
AUTHORS Actinopterygii; Neoperygii; Teleostei; Ostariophysi;			
COMMENT Cyprinidae; Cyprinidae; Danio.			
REFERENCE CK016754.1 (bases 1 to 725)			
AUTHORS NIH-MGC http://mgc.nci.nih.gov/			
COMMENT Mammalian Gene Collection (MGC)			
REFERENCE Unpublished (1999)			
AUTHORS National Institutes of Health, National Institute of Health			
COMMENT Contact: Daniela S. Gerhard, Ph.D.			
REFERENCE Office of Cancer Genomics			
AUTHORS AGO69018			
DEFINITION Pan troglodytes DNA, clone: ptB-059H17.R, genomic survey sequence.			
VERSION AGO69018.1 GI:16620820			
KEYWORDS GSS.			
ORGANISM Pan troglodytes (chimpanzee)			
KEYWORDS Pan troglodytes			
ORGANISM Bukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.			
REFERENCE 1.			
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H., and Sakaki,Y.			
TITLE BAC end sequences of Library PTB			
JOURNAL Unpublished			
FEATURES 2 (bases 1 to 917)			
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H., and Sakaki,Y.			
TITLE Direct Submission			
JOURNAL Submitted (01-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou Tsurumi-ku, Yokohama, Kanagawa 220-0045, Japan (E-mail:chimpb@gscc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: +81-45-503-9111, Fax: +81-45-503-9170)			
COMMENT Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of cloning tracking errors.			
PRIMERS Sequencing: M13REV			
VECTOR pkS145			
R.Site 1 : SacI			
R.Site 2 : SacI			
FEATURES Location/Qualifiers			
SOURCE 1..917			
ORGANISM "Pan troglodytes"			
/mol_type="genomic DNA"			
/db_xref="taxon:9598"			
/clone="PTB_059H17.R"			
/sex="male"			
/cell_type="lymphoblast"			
/clone_lib="PTB Chimpanzee Male BAC Library"			
ORIGIN			
Query Match Score 30; DB 7; Length 725;			
Best Local Similarity 44.3%; Pred. No. 26;			
Matches 31; Conservative 14; Mismatches 25; Indels 0; Gaps 0;			
Qy 2 GGAGNGGAGGAATTCAGGAGCAAUAGCAGCTCCUCUCUCAACCAGUUAAA			61
Db 663 GGAGTGATGATTCAGGAGCATGGCATGGCTGATGCTGAAAGCTATATA			604
ORIGIN			
Query Match Score 30.4%; DB 9; Length 917;			
Best Local Similarity 39.1%; Pred. No. 20;			
Matches 25; Conservative 18; Mismatches 21; Indels 0; Gaps 0;			
Qy 62 UGGGGUUGAG 71			
Db 603 TAATGATGAG 594			

RESULT 9  
 AQ108017 LOCUS  
 DEFINITION CIT-HSP-2374K6.TF CIT-HSP Homo sapiens genomic clone 2374K6,  
 Genomic survey sequence.  
 ACCESSION AQ108017  
 VERSION 1  
 KEYWORDS GSS  
 SOURCE Homo sapiens (human)  
 ORGANISM Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 ('bases 1 to 528)  
 Adams,M.D., Rounslvy,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,  
 Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and  
 Ventier,J.C.  
 TITLE Use of a random human BAC End Sequence Database for Sequence-Ready  
 Map Building  
 Unpublished (1998)  
 Other\_GSS: CIT-HSP-2374K6.TR  
 Contact: Mark Adams  
 Department of Bukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: mdreddms@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html).  
 Seq primer:: Mj3-21  
 Class: BAC ends.

FEATURES source  
 location/Qualifiers  
 1..528  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="TAXON:9606"  
 /clones="2374K6"  
 /sex="Male"  
 /cell\_type="Sperm"  
 /clone\_lib="CIT-HSP"  
 /note=Vector: pBel0BAC11; Site\_1: HindIII; Site\_2:  
 HindIII"

ORIGIN

Query Match 29.8%; Score 29.8; DB 8; Length 528;  
 Best Local Similarity 36.9%; Pred. No. 29; Gaps 0;  
 Matches 24; Conservative 19; Mismatches 22; Indels 0;  
 Locus BG838990/c

DEFINITION tae26008.y1 Hydra EST Darmstadt I mRNA linear EST 03-MAY-2004  
 similar to TR:Q9VJZ2 Q9VJZ2 CG6565 PROTEIN.; mRNA sequence.

ACCESSION CN553623  
 VERSION CN553623.1  
 KEYWORDS EST.  
 SOURCE Hydra magnipapillata  
 ORGANISM Hydridae; Hydra magnipapillata  
 Bukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;

REFERENCE 1 ('bases 1 to 578)  
 Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q., Martinez,D., Kibler,D., Hampson,S., Clifton,S., Paper,D., Narra,M.,

RESULT 10  
 CN553623/C LOCUS  
 DEFINITION tae26008.y1 Hydra EST Darmstadt I mRNA linear EST 03-MAY-2004  
 similar to TR:Q9VJZ2 Q9VJZ2 CG6565 PROTEIN.; mRNA sequence.

ACCESSION CN553623  
 VERSION CN553623.1  
 KEYWORDS EST.  
 SOURCE Hydra magnipapillata  
 ORGANISM Hydridae; Hydra magnipapillata  
 Bukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;

REFERENCE 1 ('bases 1 to 578)  
 Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q., Martinez,D., Kibler,D., Hampson,S., Clifton,S., Paper,D., Narra,M.,

COMMENT  
 Hillier,L., Martin,J., Wyllie,T., Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Ronko,I., Tsagareishvili,R., Maguire,L., Kennedy,S., Waterston,R. and Wilson,R.  
 WashU Hydra EST Project  
 Unpublished (2002)  
 Other ESTs: tae26008.x1  
 Contact: Hans Bode  
 WashU Hydra EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estowatson.wustl.edu  
 Library was constructed by Corina Gudder / GATC Konstanz, Germany  
 Library materials provided by Thomas Holstein / Molecular Cell  
 Biology, TUD, Darmstadt DNA sequencing by: Washington University  
 Genome Sequencing Center For information on obtaining a clone  
 please contact: Hans Bode (hriboe@uci.edu)  
 Seq primer: -40BP  
 High quality sequence stop: 565.  
 Location/Qualifiers  
 1..578  
 /organism="Hydra magnipapillata"  
 /mol\_type="mRNA"  
 /strain="sf-1 mutant of strain 105"  
 /db\_xref="taxon:6085"  
 /lab host="Transformax BC100 (Epicentre), T1 Phage  
 resistant electrocompetent cells"  
 /clone lib="Hydra EST Darmstadt I"  
 /note=Vector: pBluescript II SK (+); Site\_1: NotI;  
 Site\_2: EcoRI"

FEATURES source  
 location/Qualifiers  
 1..578  
 /organism="Hydra magnipapillata"  
 /mol\_type="mRNA"  
 /strain="sf-1 mutant of strain 105"  
 /db\_xref="taxon:6085"  
 /lab host="Transformax BC100 (Epicentre), T1 Phage  
 resistant electrocompetent cells"  
 /clone lib="Hydra EST Darmstadt I"  
 /note=Vector: pBluescript II SK (+); Site\_1: NotI;  
 Site\_2: EcoRI"

ORIGIN

Query Match 29.6%; Score 29.6; DB 7; Length 578;  
 Best Local Similarity 40.8%; Pred. No. 34; Gaps 0; Gaps 0;  
 Matches 31; Conservative 16; Mismatches 29; Indels 0;  
 Locus BG838990/c

DEFINITION AGGGCAUGUGGACTCGGUUCUCAACCAAGUAAAUTGGUUUAGCAAGCU 80  
 Glycine max cDNA Clone\_Gm01\_01e04, mRNA sequence.

ACCESSION 453 AAGGATACGGCACTTGTGACCATAAGACAGTGTGAATTAGTTTCATCATTTGAGT 394  
 BG838990.1 GI:14205312

VERSION EST.  
 KEYWORDS Glycine max (soybean)  
 SOURCE Glycine max  
 ORGANISM Glycine max (soybean)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicots; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

REFERENCE 1 (bases 1 to 617)  
 AUTHORS Singh,J.A., Robert,L.S., Lu,B., Zhu,L., De Moors,A., Couroux,P., Harris,L.J., Hattori,J.I., Ouellet,T., Sprott,D. and Tinker,N.A.  
 TITLE Expressed Sequence Tags from Cold-Stressed Glycine max (soybean) Leaves  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Singh,J.A.  
 Eastern Cereal and Oilseed Research Centre  
 Agriculture and Agri-food Canada  
 KW Nearby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A  
 OC, Canada  
 Tel: (613) 759-1662  
 Fax: (613) 759-1701

FEATURES Email: singhja@agr.gc.ca.  
 source Location/Qualifiers  
 1. 617 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /cultivar="Maple Arrow"  
 /db\_xref="Taxon:3847"  
 /clone="Gm01\_01e04"  
 /tissue\_type="Leaves"  
 /dev\_stage="15-day seedlings"  
 /clone\_lib="Gm01\_AAFP\_ECORC\_Glycine\_max\_cold\_stressed\_leaves"  
 /note="Vector: Bluescript SK-/xhol-ECORI; Site\_1: EcoRI;  
 Site\_2: XbaI; Plants were grown 12 days from seeds,  
 treated at 20C for 3 days, then leaves were collected.  
 Library was prepared using Unizap ZAP-CDNA synthesis kit  
 / Packaged Gigapack III Gold."

ORIGIN

Query Match 29.6%; Score 29.6; DB 7; Length 617;  
 Best Local Similarity 42.4%; Prod. No. 35; Indels 0; Gaps 0;  
 Matches 36; Conservative 14; Mismatches 35; Locus CN53289

Qy 14 UUCAUCGAGCGATAAUGUCGACTCCGTUUUCUCAAAACCAAGTTAAAUUUGGTTTATGCCA 73  
 Db 123 TGCAGGAAGCCATTGACCCACTCTCAATAAAATCAGTTAAATTGTTCCATCA 64

Query Match 29.6%; Score 29.6; DB 7; Length 617;  
 Best Local Similarity 40.8%; Prod. No. 35; Indels 0; Gaps 0;  
 Matches 31; Conservative 16; Mismatches 29; Locus CN53289

Qy 21 AGGGAAUAGGACTCCUCUCCUAAACCUAAGGUAAAATGGTTTUTGCAUAUGGCCU 80  
 Db 377 AGGGATATGGACTTGTGACCCAAATGAACTGTAAGTTGAACTAGTTCATTCAGT 436

RESULT 13

Query Match 29.6%; Score 29.6; DB 7; Length 617;  
 Best Local Similarity 40.8%; Prod. No. 35; Indels 0; Gaps 0;  
 Matches 31; Conservative 16; Mismatches 29; Locus CN53289

Qy 81 UAGGCCAACCAAGCU 96  
 Db 437 CAACAGATCAAGTT 452

ORIGIN

Query Match 29.6%; Score 29.6; DB 7; Length 617;  
 Best Local Similarity 42.4%; Prod. No. 35; Indels 0; Gaps 0;  
 Matches 36; Conservative 14; Mismatches 35; Locus CN53289

Qy 14 UUCAUCGAGCGATAAUGUCGACTCCGTUUUCUCAAAACCAAGTTAAAUUUGGTTTATGCCA 73  
 Db 123 TGCAGGAAGCCATTGACCCACTCTCAATAAAATCAGTTAAATTGTTCCATCA 64

Query Match 29.6%; Score 29.6; DB 7; Length 617;  
 Best Local Similarity 40.8%; Prod. No. 35; Indels 0; Gaps 0;  
 Matches 31; Conservative 16; Mismatches 29; Locus CN53289

Qy 21 AGGGAAUAGGACTCCUCUCCUAAACCUAAGGUAAAATGGTTTUTGCAUAUGGCCU 80  
 Db 377 AGGGATATGGACTTGTGACCCAAATGAACTGTAAGTTGAACTAGTTCATTCAGT 436

ORIGIN

Query Match 29.6%; Score 29.6; DB 7; Length 617;  
 Best Local Similarity 40.8%; Prod. No. 35; Indels 0; Gaps 0;  
 Matches 31; Conservative 16; Mismatches 29; Locus CN53289

Qy 21 AGGGAAUAGGACTCCUCUCCUAAACCUAAGGUAAAATGGTTTUTGCAUAUGGCCU 80  
 Db 466 AGGGATACGGGACTTGTGACCCAAATGAACTGTAAGTTGAACTAGTTCATTCAGT 525

REFERENCE /clone lib="Hydra EST Darmstadt I"  
 AUTHORS /note="Vector: pBluescript II SK (+); Site\_1: NotI;"  
 DEFINITION Site\_2: EcoRI"

ACCESION ta26918.x1

VERSION GI:4692593

KEYWORDS BST.

SOURCE Hydra magnipapillata

ORGANISM Eukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;

HYDRAE: Hydra

REFERENCE /base 1 to 717

AUTHORS Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q., Martinez,D., Kibler,D., Hampton,S., Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wyllie,T., Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Ronko,I., Tsagareishvili,R., Maguire,L., Kennedy,S., Waterston,R. and Wilson,R.

VERSION Unpublished (2002)

TITLE WashU Hydra EST Project

JOURNAL Contact: Hans Bode

COMMENT WashU Hydra EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

LIBRARY was constructed by Corina Guder / GATC Konstanz, Germany

LIBRARY materials provided by Thomas Holstein / Molecular Cell Biology, TUD, Darmstadt DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone  
 please contact: Hans Bode (hrbode@uci.edu)  
 Seq primer: degenerate primer.

FEATURES Location/Qualifiers

source 1. -717  
 /organism="Hydra magnipapillata"  
 /mol\_type="mRNA"

REFERENCE /strain="st-1 mutant of strain 105"

AUTHORS /db\_xref="Taxon:6085"  
 DEFINITION /lab\_host="Transformax EC100 (Epicentre), T1 Phage resistant electrocompetent cells"

LIBRARY /clone lib="Hydra EST Darmstadt I"  
 /note="Vector: pBluescript II SK (+); Site\_1: NotI;"  
 Site\_2: EcoRI"

ORIGIN

Query Match 29.6%; Score 29.6; DB 7; Length 717;  
 Best Local Similarity 40.8%; Prod. No. 36; Indels 0; Gaps 0;

Qy 21 AGGGAAUAGGACTCCUCUCCUAAACCUAAGGUAAAATGGTTTUTGCAUAUGGCCU 80  
 Db 466 AGGGATACGGGACTTGTGACCCAAATGAACTGTAAGTTGAACTAGTTCATTCAGT 525

REFERENCE /clone lib="Hydra magnipapillata"  
 AUTHORS /note="Vector: pBluescript II SK (+); Site\_1: NotI;"  
 DEFINITION Site\_2: EcoRI"

ACCESION ta26918.x1

VERSION GI:4692593

KEYWORDS BST.

SOURCE Hydra magnipapillata

ORGANISM Eukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;

HYDRAE: Hydra

REFERENCE /base 1 to 717

AUTHORS Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q., Martinez,D., Kibler,D., Hampton,S., Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wyllie,T., Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Ronko,I., Tsagareishvili,R., Maguire,L., Kennedy,S., Waterston,R. and Wilson,R.

VERSION Unpublished (2002)

TITLE WashU Hydra EST Project

JOURNAL Contact: Hans Bode

COMMENT WashU Hydra EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

LIBRARY was constructed by Corina Guder / GATC Konstanz, Germany

LIBRARY materials provided by Thomas Holstein / Molecular Cell Biology, TUD, Darmstadt DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone  
 please contact: Hans Bode (hrbode@uci.edu)  
 Seq primer: degenerate primer.

FEATURES Location/Qualifiers

source 1. -626  
 /organism="Hydra magnipapillata"  
 /mol\_type="mRNA"

REFERENCE /strain="st-1 mutant of strain 105"

AUTHORS /db\_xref="Taxon:6085"  
 DEFINITION /lab\_host="Transformax EC100 (Epicentre), T1 Phage

LIBRARY /clone lib="Hydra EST Darmstadt I"  
 /note="Vector: pBluescript II SK (+); Site\_1: NotI;"  
 Site\_2: EcoRI"

ORIGIN

Query Match 29.6%; Score 29.6; DB 7; Length 717;  
 Best Local Similarity 40.8%; Prod. No. 36; Indels 0; Gaps 0;

Qy 21 AGGGAAUAGGACTCCUCUCCUAAACCUAAGGUAAAATGGTTTUTGCAUAUGGCCU 80  
 Db 466 AGGGATACGGGACTTGTGACCCAAATGAACTGTAAGTTGAACTAGTTCATTCAGT 525

Qy 81 UAGGACAGCAACUU 96  
 Db 526 CAACGAGTCAAGTT 541

RESULT 14  
 LOCUS CN226782 744 bp mRNA Gallus gallus CDNA 5', mRNA sequence.  
 DEFINITION RJB04H01\_ab1\_RJtestis Gallus gallus EST 09-APR-2004  
 VERSION CN226782.1 GI:46330526  
 KEYWORDS EST  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.

REFERENCE 1. (bases 1 to 744)  
 AUTHORS Savolainen, P., Fitzsimmons, C.J., Arvestad, L., Andersson, L. and  
 Lundberg, J.  
 TITLE EST analysis of brain and testis cDNA libraries from White leghorn  
 and Red Jungle fowl  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Peter Savolainen  
 Department of Biotechnology  
 Royal Institute of Technology, KTH  
 SE-106 91 Stockholm, SWEDEN  
 Tel: +46 (0) 8 5537 8481  
 Fax: +46 (0) 8 5537 8335  
 Email: Peter.Savolainen@biotech.kth.se  
 Seq primer: M13 reverse primer.  
 FEATURES source  
 location/Qualifiers 1. .744  
 /organism="Gallus gallus"  
 /mol type="mRNA"  
 /strain="red junglefowl"  
 /db\_xref="taxon:9031"  
 /sex="male"  
 /lab\_host="ElectromAX DH10B (Invitrogen)"  
 /clone\_lab="Rjtestis"  
 /note="Organ: testis; Vector: PSSPORT-1; Site 1: Hind III;  
 Site 2: EcoRI; The cDNA libraries were created with the  
 Superscript Plasmid System (Invitrogen)."

ORIGIN Query Match 29.6%; Score 29.6; DB 7; Length 745;  
 Best Local Similarity 55.8%; Pred. No. 36;  
 Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;  
 Qy 1 GGAGUGGAGGAATTCAGGGAAUAGGGACUCCGCUCCUCUCAACC 52  
 Db 491 GGAGGTAAAGCTTATCGGGAAATTGGAGTCGGTCCCTCCATGCC 542

RESULT 15  
 LOCUS CO421969 745 bp mRNA Gallus gallus CDNA clone GB2ZHT1005A10, mRNA  
 DEFINITION sequence.  
 CO421969  
 ACCESSION CO421969  
 VERSION EST  
 KEYWORDS EST  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1. (bases 1 to 745)  
 AUTHORS Silve,C.S., Jorge,E.C., Patrício,M., Ledur,M.C. and Coutinho,L.L.  
 TITLE Discovery of new genes expressed in the chicken pituitary and  
 hypothalamus  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Clarissa S. Silva



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: November 30, 2004, 11:12:15 ; Search time 1093.89 Seconds

(without alignments)  
4323.081 Million cell updates/sec

Title: US-09-529-397C-25

Perfect score: 100

Sequence: 1 9ggagggaggaaatcaucagg.....uaggcagacgcaaggccucugc 100

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters:

9053458

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEntB1:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_dat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vn:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32.2	32.2	199415	10 AC084053	Mus muscu
c 2	31.2	31.2	201197	10 AC125279	Mus muscu
c 3	30.8	30.8	187605	2 AC10847	Rattus no
c 4	30.8	30.8	221981	2 AC103165	Rattus no
c 5	30.8	30.8	231814	2 AC114165	Rattus no
c 6	30.8	30.8	233823	2 AC110697	Rattus no
c 7	30.8	30.8	238172	2 AC099144	Rattus no
c 8	30.8	30.8	250740	2 AC129051	Rattus no
c 9	30.6	30.6	748	6 CQ451733	Sequence
c 10	30.6	30.6	15776	6 CQ363795	Sequence
c 11	30.6	30.6	110000	1 AE017283_15	Continuation (16 o)
c 12	30.4	30.4	21	4 AF522912	Tachylos
c 13	30.4	30.4	217	4 AF522914	Tachylos
c 14	30.4	30.4	217	4 AF522916	Tachylos
c 15	30.4	30.4	366	6 AR318058	Sequence
c 16	30.4	30.4	3185	4 AT194920	Tachylos
c 17	30.4	30.4	46077	2 AC137289	Rattus no
c 18	30.4	30.4	58143	2 AC137177	Rattus no
c 19	30.4	30.4	207260	2 AC134220	Rattus no

#### ALIGNMENTS

20	30.4	213033	2 AC131482	Rattus no
c 21	30.4	218157	2 AC139888	Rattus no
c 22	30.4	272400	2 AC110146	Rattus no
c 23	30.4	283299	2 AC112576	Rattus no
c 24	30.2	3012	2 CR536001	Ornithorh
c 25	30.2	229522	2 AC109911	Rattus no
c 26	30.2	23537	2 AC131223	Rattus no
c 27	30.2	240755	2 AC095368	Rattus no
c 28	30.0	213257	2 AC133102	Mus muscu
c 29	29.8	2938	6 AR372457	Sequence
c 30	29.8	4285	6 AR372457	Sequence
c 31	29.8	2918	8 AF057108	Populus b
c 32	29.8	2918	9 AC103805	Homo sapi
c 33	29.8	191844	2 AC148836	Pan trogl
c 34	29.8	215994	9 AC067941	Homo sapi
c 35	29.8	223143	2 CR450730	Danio rer
c 36	29.6	2339	2 AC018216	Drosophil
c 37	29.6	50481	2 AC117541	Mus muscu
c 38	29.6	79826	3 AC002443	Drosophil
c 39	29.6	110000	2 AC110832_3	Continuation (4 of 5 of
c 40	29.6	110000	2 AC11373_04	Continuation (5 of 5 of
c 41	29.6	156806	3 AC093047	Drosophil
c 42	29.6	182816	5 AC145913	Gallus ga
c 43	29.6	192132	3 AC093440	Drosophil
c 44	29.6	234627	2 AC106118	Rattus no
c 45	29.6	237005	2 AC137343	Rattus no

COMMENT  
 On Sep 5, 2002 this sequence version replaced gi:22002115.  
 Center: Harvard Partners Genome Center  
 Center Code: HPGC  
 Web site: <http://www.hpcgg.org/Sequence/mouse.html>  
 Contact: hpgc@mende1.mgh.harvard.edu

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features listing.

ANNOTATION OF FEATURES:  
 STSs are identified using ePCR (Genome Res. 7:541-550). Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double stranded sequence for all regions. All sequence is completed to a standard of coverage with a minimum of 3 reads with no ambiguities. If the sequence coverage for a region does not meet this standard, it is indicated in the annotation as Low Coverage. Low coverage linkages are verified by PCR product size verification or verification of forward and reverse reads from clones which span the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated average error rate is less than 1 per 10,000 bases using the Consed quality parameters. Regions that do not meet this requirement are annotated as Low Quality.

-----Summary Statistics  
 Center project name: ADW  
 Sequencing vector: pUC18; L08752  
 Chemistry: Dye-Terminator Big Dye; 100\*  
 Assembly program: Phrap version 0.990319  
 Contig length: 199415  
 Fraction of Phrap value < 40: 0.000326  
 Error Rate in Consed: 0.00 per 10,000 bases  
 Number of N's in consensus: 0

----- Distribution of Quality < 40 Bases:

# bases	5	10	15	20	25	30	35	40
1000	*	*	*	*	*	*	*	*
900								
800								
700								
600								
500								
400								
300								
200								
100								
0								

----- Location/Qualifiers

FEATURES source	Location/Qualifiers
	1..199415
	/organism="Mus musculus"
	/mol_type="Genomic DNA"
	/strain="C57BL/6J"
	/db_xref="Taxon:10090"
	/chromosome="3"
	/clone="RP23-158P1"
	repeat_region complement(557..1289)

repeat\_region 23882 .24118 /rpt\_family="Lx4" repeat\_region complement(24128-.24641) /rpt\_family="Lx6" repeat\_region complement(24811-.25092) /rpt\_family="Lx3" repeat\_region 25183 .25208 /rpt\_family="ITAAAn" repeat\_region 25210 .25277 /rpt\_family="GAAA)n" complement(25288-.25661) /rpt\_family="MTC" complement(26114-.32348) /rpt\_family="L1P"

Query Match 32.2% Score 32.2; DB 10; Length 199415; Best Local Similarity 49.1%; Pred. No. 8; 5; Matches 26; Conservative 14; Mismatches 13; Indels 0; Gaps 0; Features Source

37 CGGUCUNCTCUAACAGGUADAAUTGGGTTTACGCAUAUGCCUTAGGCAGAG 89  
Db 63221 CTTATCCCTAAACCAGTTAAAAATTGCTAGATATGGTAATGTGTG 63273

RESULT 2 AC125279/c LOCUS AC125279 Mus musculus BAC clone RP23-433FS from 8, complete sequence. DEFINITION ROD 05-NOV-2003 AC125279.1 GI:21536173 VERSION AC125279.1 HTG. KEYWORDS Mus musculus (house mouse) ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. REFERENCE 1. (bases 1 to 201177) AUTHORS Trani,L. and Cotton,M. TITLE Unpublished (2001) JOURNAL Submitted (21-JUN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA REFERENCES Wilson,R. TITLE Sequencing of *Mus musculus* JOURNAL Unpublished (2001) REFERENCES 2. (bases 1 to 201137) AUTHORS McPherson,J.D. and Waterston,R.H. TITLE Direct Submission JOURNAL Submitted (21-JUN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA COMMENTS Center project name: M\_BA0433FS05

MAPPING INFORMATION: Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu

SOURCE INFORMATION: The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa and Minako Tateno in the laboratory of Pieter de Jong (http://www.chori.org) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION: This sequence is the entire insert of the clone. Location/Qualifiers

1. 201197 "Mus musculus" /organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="18"  
/map=""  
/clone="RPCI-23"  
/clone\_id="RPCI-23"  
171 .557 /rpt\_family="MaLR"  
558 .1352 /rpt\_family="MaLR"  
repeat\_region 1353 .1499 /rpt\_family="Alu"  
repeat\_region 1500 .1886 /rpt\_family="MaLR"  
repeat\_region 1887 .2101 /rpt\_family="MaLR"  
repeat\_region 2103 .2485 /rpt\_family="L1"  
repeat\_region 2792 .2918 /rpt\_family="Alu"  
repeat\_region 2916 .3397 /rpt\_family="Alu"  
repeat\_region 5516 .5712 /rpt\_family="MaLR"  
repeat\_region 601 .6162 /rpt\_family="ERV"  
repeat\_region 6163 .6556 /rpt\_family="MaLR"  
repeat\_region 6557 .6980 /rpt\_family="ERV"  
repeat\_region 7137 .7220 /rpt\_family="MIR"  
repeat\_region 7970 .8262 /rpt\_family="MIR1\_type"  
repeat\_region 10573 .10940 /rpt\_family="L1"  
repeat\_region 10673 .11051 /rpt\_family="L1"  
repeat\_region 11200 .11240 /rpt\_family="L1"  
repeat\_region 12674 .12714 /rpt\_family="L1"  
repeat\_region 12715 .12907 /rpt\_family="B2"  
repeat\_region 12998 .13742 /rpt\_family="L1"  
repeat\_region 13000 .13742 /rpt\_family="L1"

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by



Shetty,J., Shvartsbeyn,A., Sission,I., Sitter,C.D., Smajs,D., Sreed,A., Soderzen,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,G., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejo,Z., Usmani,K., Valas,R., Vera,Y., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P., Williams,G., Wilson,R., Wiegert,R., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niedernsheim,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Weiss,R.A.

**TITLE** Unpublished  
**REFERENCE** 2 (bases 1 to 187606)  
**AUTHORS** Worley,K.C.  
**JOURNAL** Direct Submission  
**COMMENT** Submitted (116-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza Houston, TX 77030, USA  
**FEATURES** Direct Submission  
**TITLE** Direct Submission  
**JOURNAL** Submitted (115-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza Houston, TX 77030, USA  
**REFERENCE** 3 (bases 1 to 187606)  
**AUTHORS** Rat Genome Sequencing Consortium.  
**JOURNAL** Submitted (115-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
**COMMENT** On Nov 15, 2002 this sequence version replaced gi:23321762. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a "contig-scaffold"). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
 Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GRXS

Center Clone name: CH230-42M10

----- Summary Statistics

Assembly program: Phrap; Version 0.990329

Consensus quality: 118671 bases at least Q40

Consensus quality: 124131 bases at least Q30

Consensus quality: 127839 bases at least Q20

Estimated insert size: 122165; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length.

\* (See [http://www.hgsc.bcm.tmc.edu/docs/Genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank/draft_data.html)).

\* NOTE: This is a "working draft" sequence. It currently

\* consists of 13 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sites of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 31167: contig of 31167 bp in length

\* 2 31267: gap of unknown length

\* 31268: contig of 46116 bp in length

\* 77383: gap of unknown length

\* 77384: contig of 91428 bp in length

\* 77484: 168911: gap of unknown length

\* 168912: contig of 1410 bp in length

\* 170422: 170521: gap of unknown length

\* 170522: contig of 1508 bp in length

\* 172030 17229: gap of unknown length

\* 172130 17307: contig of 1578 bp in length

\* 17308 17307: gap of unknown length

\* 17388 17504: contig of 1297 bp in length

\* 175105 17504: gap of unknown length

\* 175205 17656: contig of 1562 bp in length

\* 176767 17888: contig of 2122 bp in length

\* 17889 17988: gap of unknown length

\* 17989 18030: contig of 1342 bp in length

\* 18031 18030: gap of unknown length

\* 18031 182073: contig of 1543 bp in length

\* 182074 182173: gap of unknown length

\* 182174 184491: contig of 2318 bp in length

\* 184492 184591: gap of unknown length

\* 184592 18706: contig of 3015 bp in length

----- Location/Qualifiers

1. 18706 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-42M10"

misc\_feature 13985..15402 /note="wgs contig"

misc\_feature 28318..29829 /note="wgs contig"

misc\_feature 77484..79244 /note="wgs contig"

misc\_feature 95087..96968 /note="wgs contig"

misc\_feature 131603..13254 /note="wgs contig"

misc\_feature 14131..15091 /note="wgs contig"

misc\_feature 16711..168911 /note="wgs contig"

----- ORIGIN

Query	Match	Score	DB	Length
Matches	Best Local Similarity	30.8%	2	187606;
Matches	Conservative	34.7%	22	Missmatches
Qy	2 GGAGGAGGAAGAUCAUCGGGAAUAGUGACUCCUCUAAACCGGUUAAGA 61			
Db	164680 GTAGGACGAGATAATTAGGAGTTATTAATCTGTGTTTCTTAATGTCATTAT 164739			
Qy	62 UGGGUUAGAUAGGCCUAGCAGAACGUCUG 99			
Db	164740 TTAGTTTGCCCTTGCCCTTGAGTTTCAGTTTCAG 164777			

----- RESULT 4

ACI03165/C	ACI03165/C	ACI03165/C	ACI03165/C
LOCUS	DEFINITION	DEFINITION	DEFINITION
Rattus norvegicus clone	Rattus norvegicus clone	CH230-194N19,	CH230-194N19,
ACI03165	VERSION	ACI03165	VERSION
GI:31578614	KEYWORDS	HTGS, PHASE2, HTGS DRAFT, HTGS FULLTOP.	KEYWORDS
HTGS, PHASE2, HTGS DRAFT, HTGS FULLTOP.	SOURCE	Rattus norvegicus (Norway rat)	SOURCE
HTGS, PHASE2, HTGS DRAFT, HTGS FULLTOP.	ORGANISM	Rattus norvegicus	ORGANISM
HTGS, PHASE2, HTGS DRAFT, HTGS FULLTOP.	MATERIAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	MATERIAL
HTGS, PHASE2, HTGS DRAFT, HTGS FULLTOP.	REFERENCE	1 (bases 1 to 21981)	REFERENCE
HTGS, PHASE2, HTGS DRAFT, HTGS FULLTOP.	AUTHORS	Muzny,D.Marie, Metzker,M.Lee, Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alisbrooks,S., Amin,A., Anguiano,D., Arnaud,J., Ayalewbechi,Y., Ayagi,A., Ayodeji,M., Baca,E., Baldwin,D., Bandaranaike,D., Barber,M., Barnesad,M., Benahmed,F., Bieswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,X., Chavez,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,	AUTHORS
HTGS, PHASE2, HTGS DRAFT, HTGS FULLTOP.	COMMENT	Muzny,D.Marie, Metzker,M.Lee, Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alisbrooks,S., Amin,A., Anguiano,D., Arnaud,J., Ayalewbechi,Y., Ayagi,A., Ayodeji,M., Baca,E., Baldwin,D., Bandaranaike,D., Barber,M., Barnesad,M., Benahmed,F., Bieswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,X., Chavez,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,	COMMENT

Center project name: GJNH  
 Center clone name: CH230-194N19  
 ----- Summary Statistics  
 Assembly program: Atlas 3.0;  
 Consensus quality: 21015 bases at least Q40  
 Consensus quality: 211432 bases at least Q30  
 Consensus quality: 216741 bases at least Q20  
 Estimated insert size: 222472; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Gebank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Gebank/draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 1 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* the accession number will be 221981 bp in length.  
 -----  
 FEATURES  
 source  
 /organism="Rattus norvegicus"  
 /mol\_type="Genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone\_id="CH230-194N19"  
 misc\_feature  
 1..1454  
 /note="wgs\_end\_extension  
 clone\_end:sp5"  
 misc\_feature  
 39859..40695  
 /note="clone\_boundary  
 clone\_end:sp6"  
 site:EcORI  
 end\_sequence:BH339559"  
 ORIGIN  
 Query Match 30.8%; Score 30.8%; Length 221981;  
 Best Local Similarity 37.8%; Pred. No. 26;  
 Matches 28; Conservative 19; Mismatches 27; Indels 0; Gaps 0;  
 Qy 26 UAUUGACUCGGCUUCCUAAACAGUUAAAATGGGTTTACCTGTATAAGGTCTAGCATGGATTAGCG 85  
 Db 43994 TTGGGGAACTCTGTTCTTCCACTACCTGTATAAGGTCTAGCATGGATTAGCG 43935  
 RESULT 5  
 AC114165/C  
 LOCUS Rattus norvegicus clone CH230-165B14, \*\*\* SEQUENCING IN PROGRESS  
 ACCESSION AC114165  
 VERSION GI:31079575  
 KEYWORDS HTGS PHASE2; HTGS DRAFT; HTGS ENRICHED.  
 SOURCE Rattus norvegicus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus;  
 REFERENCE 1 (bases 1 to 231814)  
 AUTHORS Muzny,D., Metzger,M., Lee., Abramzon,S., Adams,C., Alder,J.,  
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,  
 Anyaleibechi,V., Ayodeji,M., Baca,E., Baden,H.,  
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M.,  
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,  
 Bryant,N., Buhay,C., Burch,P., Burrel,K., Calderon,E.,  
 Cardenas,V., Carrer,K., Cavaزو,I., Caesar,H., Center,A.,  
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chu,J.,  
 -----  
 JOURNAL Genome Center  
 AUTHORS Unpublished  
 TITLE Direct Submission  
 JOURNAL Submitted (24-NOV-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 221981)  
 REFERENCE Rat Genome Sequencing Consortium.  
 JOURNAL Submitted (13-MAY-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On May 13, 2003 this sequence was replaced gi:2507265.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.  
 -----  
 JOURNAL Genome Center  
 CENTER Baylor College of Medicine  
 CENTER CODE BCM  
 CENTER WEB <http://www.hgsc.bcm.tmc.edu>  
 CONTACT hgsc-help@bcm.tmc.edu  
 ----- Project Information

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denison,S., Deramo,C., Dunn,A., Durbin,K., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,B., Eaves,R., Egan,A., Escottto,M., Eugene,C.A., Evans,C.A., Fallis,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Gantara,R., Garcia,A., Garner,T., Garza,M., Gebregesorgis,E., Geer,K., Gil,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamilton,C., Hamilton,K., Harvey,Y., Havlik,P., Hayes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Gladun,S.L., Hogson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levant,J., Lewis,S., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Loresuhera,L., Louisaged,H., Lozano R.J., Lu,X., Ma,J., Maheshwari,M., Mahmud,M., Mallory,K., Mangum,A., Mangun,B., Mapua,P., Martin,K., Martinez,E., Mawhinney,S., Mcleod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montenayor,J.J., Moore,S., Morgan,M., Morris,M., Neal,D., Newton,N., Nguyen,N., Norris,S., Nair,L., Nankervis,C., Okwonou,G., Olarmannsson,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Ross,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sned,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villanova,L., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wieczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakuob,S., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G., and Gibbs,R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 231814)

TITLE JOURNAL  
AUTHORS  
TITLE JOURNAL  
REFERENCE JOURNAL  
AUTHORS  
TITLE JOURNAL  
COMMENT JOURNAL  
DEFINITION LOCUS  
ACCESSION AC110697  
KEYWORDS HTGS PHASE2; HTGS\_DRAFT; HTGS\_ENRICHED.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

----- Project Information  
Center Project name: GIOV  
Center Clone name: CH230-165B14  
Summary Statistics  
Assembly Program: Atlas 3.0;  
Consensus Quality: 219004 bases at least Q40  
Consensus Quality: 220024 bases at least Q30  
Consensus Quality: 224099 bases at least Q20  
Estimated Insert size: 229718; sum-of-contigs estimation  
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.biosc.bcm.edu/docs/Genbank/draft\\_data.html](http://www.biosc.bcm.edu/docs/Genbank/draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* are believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 231814: contig of 231814 bp in length.  
1. 231814  
FEATURES  
Source  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone\_id="CH230-165B14"  
1..1811  
/note="wgs\_contig"  
1  
/note="clone boundary  
clone\_end:SP6  
site:ECRI  
end sequence:BH272064"  
93963..96094  
/note="wgs contig"  
119152..121641  
/note="wgs contig"  
/note="clone boundary  
clone\_end:R7  
site:ECRI  
end sequence:BH272063"  
223398..223814  
/note="wgs end extension  
clone\_end:T7"  
ORIGIN  
Query Match 30.8%; Score 30.8; DB 2; Length 231814;  
Best Local Similarity 34.7%; Pred. No. 26;  
Matches 34; Conservative 22; Mismatches 42; Indels 0; Gaps 0;  
Qy 2 GGAGGGAGAAAUCAUAGGGRAGAUAGCUCGCTTCCUCAAACGAGUUA 61  
Db 230424 GTAGGGCAGATAATGAGGTATTAAATCTGTTTCCTAAGTCATTAT 230365  
Qy 62 UGGGUUUUAGCAUAGGCCUTAGGCCAGGAAGCTAGTTAGTCAG 99  
Db 230364 TTAGTTTGCCCTAGTCCTGGCCATAGTTAGTCAG 230327  
RESULT 6  
AC110697  
DEFINITION Rattus norvegicus clone CH230-43H7  
ACCESSION AC110697  
KEYWORDS HTGS PHASE2; HTGS\_DRAFT; HTGS\_ENRICHED.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

----- Genome Center  
Center code: BCM  
Web site: <http://www.biosc.bcm.edu>  
Contact: hgsc-help@bcm.tmc.edu



RESULT	7	JOURNAL	Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
AC099144/c	AC099144	COMMENT	On May 10, 2003 this sequence version replaced g1:2326459.
LOCUS	Rattus norvegicus clone CH230-68F7, WORKING DRAFT SEQUENCE.	DEFINITION	The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas ( <a href="http://www.hgsc.bcm.edu/projects/rat/">http://www.hgsc.bcm.edu/projects/rat/</a> ). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
ACCESSION	AC099144	ORGANISM	Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus norvegicus
VERSION	AC099144.5	KEYWORDS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus.
AUTHORS	Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Aisbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoysgi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaraanake, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blith, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesari, H., Center, A., Chacko, J., Chavez, D., Chan, G., Chen, R., Chen, Y., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Divila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escott, M., Eugene, C., Evans, C.A., Fallis, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisa, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Geurts, E., Geer, K., Gill, R., Grady, M., Guerra, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hernandez, J., Harvey, Y., Havlik, P., Hayes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howell, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levyan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewala, L., Louisado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindarne M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapka, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meinen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakolemeh, O., Okwonou, G., Olarnpusungsaoon, A., Pal, S., Parks, K., Pasternak, S., Perez, A., Perez, L., Perez, P., Pfannkoch, C., Plopper, F., Poundexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Ross, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Snead, A., Sodergren, B., Song, X.-Z., Sorrelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, P., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejo, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wlecyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Niederauer, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinrock, G., and Gibbs, R.A.		
REFERENCE	1 (bases 1 to 238172)	JOURNAL	Unpublished
AUTHORS	Worley, K.C.	COMMENT	Direct Submission
TITLE	AC099144	DEFINITION	Submitted (09-NCV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
KEYWORDS	3 (bases 1 to 238172)	ORGANISM	Rat Genome Sequencing Consortium.
SOURCE		REFERENCE	3 (bases 1 to 238172)
ORGANISM		AUTHORS	Direct Submission
		TITLE	23 GCAU AUGUGACTCCGCTCTTAAACAGTATAAATGGUTUUGCAUAGGCCU 80



misc\_feature 179234..181122  
 /note="wgs\_contig"

ORIGIN

Query Match 30.6%; Score 30.6; DB 2; Length 250740;  
 Best Local Similarity 43.1%; Pred. No. 26;  
 Matches 25; Conservative 16; Mismatches 17; Indels 0; Gaps 0;

Qy 23 GCAUAUGGAGTCGGTCTTCAAAACCAAGTAAATGGTTTGGAUAGGCCU 80  
 Db 4082 GCATATTGATTCCTTGTCCCTAACCTTGAGCTGTACAATCTC 4025

RESULT 9  
 CQ451733/C CQ451733 748 bp DNA linear PAT 30-JAN-2004  
 DEFINITION Sequence 17493 from Patent WO0192523.  
 VERSION CQ451733.1 GI:41420097  
 KEYWORDS  
 ORGANISM Homo sapiens (human)

Mammalia; Eutheria; Primates; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Metazoa; Catarrhini; Hominidae; Homo.

REFERENCE 1. Shimkets, R.A. and Leach, M.D.  
 Human polynucleotides and polypeptides encoded thereby  
 Patent: WO 019523-A 17/93 06-DEC-2001;  
 Curagen Corporation (US)

FEATURES Location/Qualifiers  
 1..748  
 source /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Query Match 30.6%; Score 30.6; DB 6; Length 748;  
 Best Local Similarity 46.8%; Pred. No. 20;  
 Matches 36; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

Qy 1.8 UCGAGGCAUAGGAGCACCGGUCCUCUCAAACCAGTAAATGGTTTGGAUAGGCCU 77  
 Db 317 TRGGGTTGAGCTTGACCCAGGTTCTTCAGACCGTTGARGAAACGGTTAGGTCTG 258

Qy 78 CCTUAGGAGCAGGAAGC 94  
 Db 257 CCCGGGGAGCGGAAGC 241

Query Match 30.6%; Score 30.6; DB 6; Length 15776;  
 Best Local Similarity 46.8%; Pred. No. 25;  
 Matches 36; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

Qy 18 UCGAGGCAUAGGAGCACCGGUCCUCUCAAACCAGTAAATGGTTTGGAUAGGCCU 77  
 Db 8427 TTGGCTGAGCTCAGCGCTCTGAGCGTTGAGCTGATGAACTGTTAGGTCTG 8486

Qy 78 CCTUAGGAGCAGGAAGC 94  
 Db 8487 CCCGGACGGAAAGC 8503

RESULT 11  
 AE017283\_15  
 WPCOMMIT

Sequence split into 26 fragments LOCUS AE017283 Accession AE017283  
 Fragment Name Begin End

AE017283\_00 100000 110000  
 AE017283\_01 200000 310000  
 AE017283\_02 300000 410000  
 AE017283\_03 400000 510000  
 AE017283\_04 500000 610000  
 AE017283\_05 600000 710000  
 AE017283\_06 700000 810000  
 AE017283\_07 800000 910000  
 AE017283\_08 900000 1010000  
 AE017283\_09 1000000 1110000  
 AE017283\_10 1100000 1210000  
 AE017283\_11 1200000 1310000  
 AE017283\_12 1300000 1410000  
 AE017283\_13 1400000 1510000  
 AE017283\_14 1500000 1610000  
 AE017283\_15 1600000 1710000  
 AE017283\_16 1700000 1810000  
 AE017283\_17 1800000 1910000  
 AE017283\_18 1900000 2010000  
 AE017283\_19 2000000 2110000  
 AE017283\_20 2100000 2210000  
 AE017283\_21 2200000 2310000  
 AE017283\_22 2300000 2410000  
 AE017283\_23 2400000 2510000  
 AE017283\_24 2500000 2560265  
 Continuation T<sub>16</sub> of 26) of AE017283 from base 1500001 (AE017283 Propionibacterium acnes F

Query Match 30.6%; Score 30.6%; DB 1; Length 110000;  
 Best Local Similarity 46.8%; Pred. No. 29;  
 Matches 36; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

Qy 18 UCGAGGCAUAGGAGCACCGGUCCUCUCAAACCAGTAAATGGTTTGGAUAGGCCU 77  
 Db 34666 TTGGCTGAGCTCAGCGCTCTGAGCGTTGAGCTGATGAACTGTTAGGTCTG 34725

Qy 78 CCTUAGGAGCAGGAAGC 94  
 Db 34726 CCCGGACGGAAAGC 34742

RESULT 12  
 AF522912/C  
 LOCUS Tachyglossus aculeatus clone Tag7 type I interferon gene, Partial  
 DEFINITION cds

ACCESSION AF522912  
 VERSION AF522912.1 GI:21451581  
 KEYWORDS Tachyglossus aculeatus (Australian echidna)  
 SOURCE Tachyglossus aculeatus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelostomi;  
 Mammalia; Monotremata; Tachyglossidae; Tachyglossus.

REFERENCE 1 (bases 1 to 217)

AUTHORS	Harrison,G.A., Young,L.J., Watson,C.M., Miska,K.B., Miller,R.D. and Deane,B.M.	FEATURES	Location/Qualifiers
TITLE	A survey of type I interferons from a marsupial and monotreme: implications for the evolution of the type I interferon gene family in mammals	source	1. .217 /organism="Tachyglossus aculeatus" /mol_type="genomic DNA" /db_xref="taxon:9261" /clone="TaG11"
JOURNAL	Cytokine 21 (3), 105-119 (2003)	mRNA	<1. .>17 /product="type I interferon" <1. .>217 /codon_start=2 /product="type I interferon" /protein_id="AA014986_1" /db_xref="GI:27451586" /translation="NIFSRNHSQTGMDEAVKEFLHGQHQMWWLFLFEEGMWNS TLRRDISHIKSYFKRMMDLKGRNYS"
MEDLINE	22583464	CDS	
PUBLMED	12697149		
REFRENCES	2 (bases 1 to 217)		
AUTHORS	Harrison,G.A., Young,L.J., Watson,C.M., Miska,K.B., Miller,R.D. and Deane,B.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-JUN-2002) School of Science, Food and Horticulture, University of Western Sydney, Locked Bag 1797, Penrith South DC, NSW 1797, Australia		
FEATURES	Location/Qualifiers		
source	1. .217 /organism="Tachyglossus aculeatus" /mol_type="genomic DNA" /db_xref="taxon:9261" /clone="TaG11"		
mRNA	<1. .>217 /product="type I interferon" <1. .>217 /codon_start=2 /product="type I interferon" /protein_id="AA014986_1" /db_xref="GI:27451586" /translation="NIFSRNHSQTGMDEAVKEFLHGQHQMWWLFLFEEGMWNS TLRRDISHIKSYFKRMMDLKGRNYS"		
CDS			
ORIGIN			
		Query Match 30.4%; Score 30.4; DB 4; Length 217;	
		Best Local Similarity 41.7%; Pred. No. 21;	
		Matches 30; Conservative 16; Mismatches 26; Indels 0; Gaps 0;	
		ACCESSION AF522916/C	
		DEFINITION Tachyglossus aculeatus clone TaG31 type I interferon gene, partial cds	
		VERSION AF522916	
		KEYWORDS	
		SOURCE	
		ORGANISM	
		REFERENCE	
		AUTHORS	
		Deane, B.M.	
		TITLE	
		A survey of type I interferons from a marsupial and monotreme: implications for the evolution of the type I interferon gene family in mammals	
		JOURNAL Cytokine 21 (3), 105-119 (2003)	
		MEDLINE 22583464	
		PUBMED 12697149	
		REFERENCE 2 (bases 1 to 217)	
		AUTHORS Harrison,G.A., Young,L.J., Watson,C.M., Miska,K.B., Miller,R.D. and Deane,B.M.	
		TITLE JOURNAL	
		Submitted (19-JUN-2002) School of Science, Food and Horticulture, University of Western Sydney, Locked Bag 1797, Penrith South DC, NSW 1797, Australia	
		FEATURES	
		source	
		1. .217 /organism="Tachyglossus aculeatus" /mol_type="genomic DNA" /db_xref="taxon:9261" /clone="TaG11"	
		/product="type I interferon" <1. .>217 /codon_start=2 /product="type I interferon" /protein_id="AA014986_1" /db_xref="GI:27451590" /translation="NIFSRNHSQTGMDEAVKEFLHGQHQMWWLFLFEEGMWNS TLRRDISHIKSYFKRMMDLKGRNYS"	
		CDS	
JOURNAL	Cytokine 21 (3), 105-119 (2003)		
MEDLINE	22583464		
PUBLMED	12697149		
REFRENCES	2 (bases 1 to 217)		
AUTHORS	Harrison,G.A., Young,L.J., Watson,C.M., Miska,K.B., Miller,R.D. and Deane,B.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-JUN-2002) School of Science, Food and Horticulture, University of Western Sydney, Locked Bag 1797, Penrith South DC, NSW 1797, Australia		

## ORIGIN

Query Match 30.4%; Score 30.4; DB 4; Length 217;  
 Best Local Similarity 41.7%; Pred. No. 21;  
 Matches 30; Conservative 16; Mismatches 26; Indels 0; Gaps 0;  
 Matches 30;保守性 16; 错配 26; 插入 0; 缺失 0;

Qy 6 UGGAGGAUUAUCUGGCAUAUGUGCAUCGGUCUCUTCAAACCGATUAUAATCG 65  
 Db 78 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 78 TGATGGACTCCATGGAGAATTCTCGAACTATGGCTCATGCCAGCTGTTGGAGTGA 19

Qy 66 UUUUAGCUAUG 77  
 Db 18 TTTCTGTCTGAAG 7

## RESULT 15

AR318058/c AR318058 366 bp DNA linear PAT 17-AUG-2003

DEFINITION Sequence 608 from patent US 652958.

ACCESSION AR318058

VERSION AR318058.1 GI:33699161

KEYWORDS

Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 366)

AUTHORS Breton, G. and Bush, D.

TITLE Nucleic acid and amino acid sequences relating to Acinetobacter

baumanni for diagnostics and therapeutics

JOURNAL Patent: US 652958-A 608 13-MAY-2003;

FEATURES location/qualifiers

source 1..366

/organism="unknown"

/mol\_type="genomic DNA"

## ORIGIN

Query Match 30.4%; Score 30.4; DB 6; Length 366;  
 Best Local Similarity 39.8%; Pred. No. 22;  
 Matches 35; Conservative 17; Mismatches 36; Indels 0; Gaps 0;

Matches 35;保守性 17; 错配 36; 插入 0; 缺失 0;

Qy 1 GGGAGUGGAGAUATACUGGCAUAGUGCAUCGGUCUCUTCAAACCGATUAUA 60

Db 101 GGTTGTTATGATAATAGCTTCTATATGTCSCACCAAGTAATTGTCAAACAAAGTAATAG 42

Qy 61 AUUGGUUUAUGCAUAUGGUUUCGACA 88

Db 41 TCTGGTTATGCTATTCAGATTCA 14

Search completed: November 30, 2004, 11:52:20  
 Job time : 1100.89 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw mode!

Run on: November 30, 2004, 11:12:11 ; Search time 232.222 Seconds  
(without alignments)

22660.516 Million cell updates/sec

Title: US-09-529-397C-25

Perfect score: 100

Sequence: 1 gggtggggggaaaucaucg.....uaggccacagcaaggccuucugc 100

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2524710521 residues

Total number of hits satisfying chosen parameters:

8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing First 45 summaries

Database : N\_Geneseq\_23Seq04:  
1: geneseqn1980s:/\*  
2: geneseqn1990s:/\*  
3: geneseqn2000s:/\*  
4: geneseqn200as:/\*  
5: geneseqn200bs:/\*  
6: geneseqn200as:/\*  
7: geneseqn200bs:/\*  
8: geneseqn200as:/\*  
9: geneseqn200bs:/\*  
10: geneseqn2003cs:/\*  
11: geneseqn2003ds:/\*  
12: geneseqn2004s:/\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	100	100.0	100	3	AAZ99048		Aaz99048 RNA aptam
2	100	100.0	100	12	ADJ10063		Adj10063 Synthetic
3	100	100.0	100	12	ADJ10064		Adj10065 Synthetic
4	100	100.0	200	12	ADJ10065		Aaz99050 RNA aptam
5	90	90.0	90	3	AAZ99050		Adj10053 RNA aptam
6	84.4	84.4	98	12	ADJ10053		Aaz99051 RNA aptam
7	80	80.0	80	3	AAZ99051		Adj10066 Synthetic
8	50.6	50.6	147	12	ADJ10066		Aaz99056 Ras target
9	45	45	45	3	AAZ99076		Aac75113 Human ORF
c 10	30.6	30.6	748	6	ABN24508		Abn24508 Propionib
c 11	30.6	30.6	15776	4	AAS59583		Acf64512 Propionib
c 12	30.6	30.6	15776	8	ACF64512		Adz23321 DNA encod
c 13	30.6	30.6	15776	9	ADA23321		Aaz99049 RNA aptam
c 14	30.4	30.4	366	9			Aaz99049 RNA aptam
c 15	30	30.0	98	3	AAZ99049		Aaz75942 Poplar fl
c 16	29.8	29.8	4192	3	AAZ95391		Aaz85391 Nucleotid
c 17	29.8	29.8	4285	4	ABK88484		Abk88484 Poplar pr
c 18	29.8	29.8	4285	6	ACA62517		Aca62517 Poplar ho
c 19	29.8	29.8	4285	9	AAZ99084		Aaz99084 Ras target
c 20	29	29	3	12	ADJ10055		Adj10055 Synthetic

ACa52761 Prokaryot  
Abz15103 Arabidops  
Aby24321 Human pro  
Aax58401 Vicia sat  
Aij10058 PCR prime  
Adj10057 PCR prime  
Adj10056 PCR prime  
Continuation (2 of 2 of  
Continuation (2 of 2 of  
Continuation (2 of 2 of  
Aai19497 Human neu  
Abt12823 Human neu  
Aat3575 Chicken b  
Aaz99087 Ras target  
Aach7425 Human sec  
Aaa69206 Bacteriop  
Aaa69168 Bacteriop  
Aac21986 Human sec  
Aah51835 HIV prote  
Adh4700 HIV Prote  
Ada2579 Mouse Rel  
Adb72317 Mouse Rel

ALIGNMENTS

RESULT 1  
AAZ99048  
ID AAZ99048 standard; RNA; 100 BP.  
XX  
AAZ99048;  
XX  
21-JUN-2000 (first entry)  
XX  
DE RNA aptamer #25 for binding Ras target protein.  
XX  
Ras target protein; malignant tumour; signal transmission regulation; cell proliferation; cell differentiation; aptamer; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200009684-A1.  
XX  
PD 24-FEB-2000.  
XX  
PP 13-AUG-1999; 99WO-JP04399.  
XX  
PR 14-AUG-1998; 98JP-00242596.  
PR 24-NOV-1998; 98JP-0033284.  
XX  
(NISC-) JAPAN SCI & TECHNOLOGY CORP.  
XX  
Yokoyama, S., Hirao I., Sakamoto K.;  
XX  
WPI; 2000-224330/19.  
DR  
XX  
PT Nucleic acid e.g. RNA aptamer capable of binding specifically to Ras target protein like Raff1, useful in drug compositions to treat and diagnose malignant tumors and inflammation.  
XX  
PS Claim 6; Page 41; 59pp; Japanese.  
XX  
CC The invention relates to novel nucleic acids which bind specifically to the target protein of Ras, e.g. Raf-1, RNA aptamer (AAZ99024-299051) based on these sequences are useful in the treatment and diagnosis of malignant tumours and inflammation. The nucleic acids can be used to formulate medicinal compositions that are useful in the treatment of malignant tumours and inflammation and for disease diagnosis by binding specifically to Ras target protein and regulating transmission of signal causing proliferation or differentiation of cells

CC exemplification of the invention.

XX Sequence 100 BP; 25 A; 22 C; 24 G; 0 T; 29 U; 0 Other;

SQ Query Match 100.0%; Score 100; DB 3; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.4e-25;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGAGUGGAGGAUTCAUGGAGGAUAGUCGUCCUCUCAACCGAUUA 60  
Db 1 GGGAGUGGAGGAUTCAUGGAGGAUAGUCGUCCUCUCAACCGAUUA 60

QY 61 AUUGGUUUTUAGCAUAUGCCUAGCAGCACGCTUCUGC 100  
Db 61 AUUGGUUUTUAGCAUAUGCCUAGCAGCACGCTUCUGC 100

RESULT 2  
ADJ10063

XX ADJ10063 standard; DNA; 100 BP.

AC ADJ10063;

XX DT 17-JUN-2004 (first entry)

DB Synthetic RNA 9A (100-mer).

XX ss; unnatural base; 5' substituted-2-oxo(1H)- Pyridin-3-yl;  
photo-reactive group; biotin; fluorescent molecule;  
6-substituted-2-aminopurin-9-yl; aptamer; Gene therapy.

OS Synthetic.

XX WO2004007713-A1.

XX PD 22-JAN-2004.

XX PP 28-FEB-2003; 2003WO-JP002342.

XX PR 17-JUL-2002; 2002JP-00208568.

XX PA (RIKE ) RIKEN KK.  
(NISC-) JAPAN SCI & TECHNOLOGY CORP.  
(ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.

PA PI Hirao I., Yokoyama S., Hirao M., Mitsui T;

DR WPI; 2004-122944-12.

PS Disclosure; Fig 10; 78PP; Japanese.

XX This invention relates to nucleosides or nucleotides that a novel unnatural base, namely the 5'-substituted-2-oxo(1H)- pyridin-3-yl group as a base. Specifically, it refers to nucleic acid molecules integrated with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as base which is applied as template for the integration of any of the nucleotides into its complementary position. Accordingly, the nucleosides and nucleotides are useful in developing functional nucleic acids including antisense DNAs and RNAs, ribozymes and aptamers as reagents or therapeutics in treating diseases by gene therapy. The produced nucleic acids are chemically stable with improved ease of amplification, replication and transcription for their preparation. This polynucleotide sequence is a synthetic RNA 9A aptamer (100-mer) given in an

CC Sequence 100 BP; 25 A; 22 C; 24 G; 0 T; 29 U; 0 Other;

SQ Query Match 100.0%; Score 100; DB 12; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.4e-25;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGAGUGGAGGAUTCAUGGAGGAUAGUCGUCCUCUCAACCGAUUA 60  
Db 1 GGGAGUGGAGGAUTCAUGGAGGAUAGUCGUCCUCUCAACCGAUUA 60

QY 61 AUUGGUUUTUAGCAUAUGCCUAGCAGCACGCTUCUGC 100  
Db 61 AUUGGUUUTUAGCAUAUGCCUAGCAGCACGCTUCUGC 100

RESULT 3  
ADJ10064

XX ID ADJ10064 standard; DNA; 100 BP.

XX AC ADJ10064;

XX DT 17-JUN-2004 (first entry)

DB Synthetic RNA 9A (5Iy87) aptamer (100-mer).

XX DE ss; unnatural base; 5' substituted-2-oxo(1H)- Pyridin-3-yl;  
photo-reactive group; biotin; fluorescent molecule;  
6-substituted-2-aminopurin-9-yl; aptamer; gene therapy.

XX OS Synthetic.

XX FH Key FT modified\_base 87  
FT /mod\_base OTHER  
FT /note= "OTHER= 5Iy"  
FT /note= "OTHER= 5Iy"

XX Location/Qualifiers

XX PR WO2004007713-A1.

XX PD 22-JAN-2004.

XX PR 17-JUL-2002; 2002JP-00208568.

XX PA (RIKE ) RIKEN KK.  
(NISC-) JAPAN SCI & TECHNOLOGY CORP.  
(ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.

PA PI Hirao I., Yokoyama S., Hirao M., Mitsui T;

XX DR WPI; 2004-122944/12.

XX PS Disclosure; Fig 10; 78PP; Japanese.

XX This invention relates to nucleosides or nucleotides comprising novel base particularly 5-substituted-2-oxo(1H)-pyridin-3-yl group as base, useful in developing functional nucleic acids as reagents or therapeutics in treating diseases.

XX Disclosure; Fig 10; 78PP; Japanese.

CC This invention relates to nucleosides or nucleotides that a novel unnatural base, namely the 5'-substituted-2-oxo(1H)- pyridin-3-yl group as a base. Specifically, it refers to nucleic acid molecules integrated with a base substituted at the 5-position with either a photo-reactive group (e.g. iodine or bromine), an alkeneyl, alkyneyl or amino group, biotin or derivative thereof or a fluorescent molecule selected from fluorescein, 6-carboxyfluorescein, tetramethyl-1,7-carboxyrhodamine or fluorescein, 6-carboxyfluorescein, tetraethyl-1,7-carboxyrhodamine or derivatives thereof. The present invention describes preparing nucleic acids by carrying out transcription, replication or reverse transcription with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as base which is applied as template for the integration of any of the nucleotides into its complementary position. Accordingly, the nucleosides and nucleotides are useful in developing functional nucleic acids including antisense DNAs and RNAs, ribozymes and aptamers as reagents or therapeutics in treating diseases by gene therapy. The produced nucleic acids are chemically stable with improved ease of amplification, replication and transcription for their preparation. This polynucleotide sequence is a synthetic RNA 9A aptamer (100-mer) given in an

CC exemplification of the invention.

CC This invention relates to nucleosides or nucleotides that a novel unnatural base, namely the 5'-substituted-2-oxo(1H)- pyridin-3-yl group as a base. Specifically, it refers to nucleic acid molecules integrated with a base substituted at the 5-position with either a photo-reactive group (e.g. iodine or bromine), an alkeneyl, alkyneyl or amino group, biotin or derivative thereof or a fluorescent molecule selected from fluorescein, 6-carboxyfluorescein, tetramethyl-1,7-carboxyrhodamine or fluorescein, 6-carboxyfluorescein, tetraethyl-1,7-carboxyrhodamine or derivatives thereof. The present invention describes preparing nucleic acids by carrying out transcription, replication or reverse transcription with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as base which is applied as template for the integration of any of the

CC nucleotides into its complementary position. Accordingly, the nucleosides CC and nucleotides are useful in developing functional nucleic acids CC including antisense DNAs and RNAs, ribozymes and aptamers as reagents or CC therapeutics in treating diseases by gene therapy. The produced nucleic CC acids are chemically stable with improved ease of amplification, CC replication and transcription for their preparation. This polynucleotide CC sequence is a synthetic RNA 9A(5Iy87) aptamer (100-mer) given in an CC exemplification of the invention.

XX Sequence 100 BP; 25 A; 22 C; 24 G; 0 T; 29 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 100;  
Best Local Similarity 100.0%; Pred. No. 2.4e-25;  
Matches 100; Conservative 0; Mismatches 0; Gaps 0;

OY 1 GGAGUGGAGGAUUCAUAGGAGCAUAUGGUCCACUCCUCUTCAACCAGUTAAA 60  
DB 1 GGAGUGGAGGAUUCAUAGGAGCAUAUGGUCCACUCCUCUTCAACCAGUTAAA 60  
QY 61 AUUGGTTTATGCAAUAGGCCUUAAGGACAGCAAGCUUCUGC 100  
DB 61 AUUGGTTTATGCAAUAGGCCUUAAGGACAGCAAGCUUCUGC 100

RESULT 4  
ADJ10065 standard; DNA; 200 BP.  
XX ADJ10065;  
XX 17-JUN-2004 (first entry)

DE Synthetic RNA 2x9A aptamer (200-mer).

XX ss; unnatural base; 5'-substituted-2'-oxo(1H)- Pyridin-3-yl;  
XX photoreactive group; biotin; fluorescent molecule;  
XX 6'-substituted-2'aminopurin-9-yl; aptamer; gene therapy.  
OS Synthetic.  
XX WO2004007713-A1.  
XX 22-JAN-2004.  
XX 28-FEB-2003; 2003WO-JP002342.  
XX 17-JUL-2002; 2002JP-00208568.

XX (RIKE ) RIKEN KK.  
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.  
XX PI Hirao I, Yokoyama S, Hirao M, Mitsui T;  
XX DR; 2004-122944/12.  
XX Disclosure: Fig 10; 78pp; Japanese.

PS Nucleosides or nucleotides comprising novel base particularly 5'-substituted-2'-oxo(1H)-pyridin-3-yl group as base, useful in developing functional nucleic acids as reagents or therapeutics in treating diseases.  
XX Disclosure: Fig 10; 78pp; Japanese.

PS This invention relates to nucleosides or nucleotides that a novel unnatural base, namely the 5'-substituted-2'-oxo(1H)- pyridin-3-yl group as a base. Specifically, it refers to nucleic acid molecules integrated with a base substituted at the 5-position with either a photoactive group (e.g. iodine or bromine), an alkenyl, alkynyl or amino group, biotin or a derivative thereof or a fluorescent molecule selected from fluorescein, 6-carboxyfluorescein, tetramethyl-7-carboxyrhodamine or derivatives thereof. The present invention describes preparing nucleic acids by carrying out transcription, replication or reverse transcription with a nucleic acid containing a 6-substituted-2-aminopurin-3-yl group as

base which is applied as template for the integration of any of the nucleotides into its complementary position. Accordingly, the nucleosides and nucleotides are useful in developing functional nucleic acids including antisense DNAs and RNAs, ribozymes and aptamers as reagents or therapeutics in treating diseases by gene therapy. The produced nucleic acids are chemically stable with improved ease of amplification, replication and transcription for their preparation. This polynucleotide sequence is a synthetic RNA 2x9A aptamer (200-mer) given in an exemplification of the invention.

XX SQ Sequence 200 BP; 50 A; 44 C; 48 G; 0 T; 58 U; 0 Other;  
SQ Sequence Match 100.0%; Score 100; DB 12; Length 200;  
Best Local Similarity 100.0%; Pred. No. 3e-25;  
Matches 100; Conservative 0; Mismatches 0; Gaps 0;  
QY 1 GGAGUGGAGGAUUCAUAGGAGCAUAUGGUCCACUCCUCUTCAACCAGUTAAA 60  
DB 1 GGAGUGGAGGAUUCAUAGGAGCAUAUGGUCCACUCCUCUTCAACCAGUTAAA 60  
QY 61 AUUGGTTTATGCAAUAGGCCUUAAGGACAGCAAGCUUCUGC 100  
DB 61 AUUGGTTTATGCAAUAGGCCUUAAGGACAGCAAGCUUCUGC 100

RESULT 5

AAZ99050  
ID AAZ99050 standard; RNA; 90 BP.  
XX AC AAZ99050;  
XX AC  
XX 21-JUN-2000 (first entry)  
XX RNA aptamer #27 for binding Ras target protein.  
XX DE RNA target protein; malignant tumour; signal transmission regulation; cell proliferation; cell differentiation; aptamer; inflammation; ss.  
XX KW Homo sapiens.  
XX OS  
XX PN WO200009884-A1.  
XX PD 24-FEBB-2000.  
XX PP 13-AUG-1999; 99WO-JP004399.  
XX PR 14-AUG-1998; 98JP-00242596.  
XX PR 24-NOV-1998; 98JP-0033284.  
XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
XX PI Yokoyama S, Hirao I, Sakamoto K;  
XX WPI; 2000-224430/19.  
XX DR  
XX Disclosure: Fig 10; 78pp; Japanese.

PS The invention relates to novel nucleic acids which bind specifically to the target protein of Ras, e.g. Raf-1. RNA aptamer (AAZ99024-299051) based on these sequences are useful in the treatment and diagnosis of malignant tumours and inflammation. The nucleic acids can be used to formulate medicinal compositions that are useful in the treatment of malignant tumours and inflammation and for disease diagnosis by binding specifically to Ras target protein and regulating transmission of signal causing proliferation or differentiation of cells.

XX SQ Sequence 90 BP; 23 A; 19 C; 22 G; 0 T; 26 U; 0 Other;  
SQ Query Match 90.0%; Score 90; DB 3; Length 90;

Qy	1	GGAGUGGAGGAUCAUCAGGAUAUCUGAACUCCGUUCUCCUAAACCGUUA	60	Query Match Score 84.4%; Best Local Similarity 97.0%; Pred. No. 7; Mismatches 0; Conservation 0; Matches 90; Gaps 0;
Db	1	GGAGUGGAGGAUCAUCAGGAUAUCUGAACUCCGUUCUCCUAAACCGUUA	60	
Qy	61	AUUGGUUTTACGCAUAGGCCCTAAGGCACAGC	90	Qy 1 GGAGUGGAGGAUCAUCAGGAUAUCUGAACUCCGUUCUCCUAAACCGUUA 60
Db	61	AUUGGUUTTACGCAUAGGCCCTAAGGCACAGC	90	Db 1 GGAGUGGAGGAUCAUCAGGAUAUCUGAACUCCGUUCUCCUAAACCGUUA 60
				Qy 61 AUUGGUUTTACGCAUAGGCCCTAAGGCACAGC 100
				Db 59 AUUGGUUTTACGCAUAGGCCCTAAGGCACAGC 98
<b>RESULT 6</b>				
ADJ10053	ID ADJ10053 standard; RNA; 98 BP.			
XX				
AC				
XX				
DT	17-JUN-2004 (first entry)			
XX				
DE	RNA aptamer (RNA9A) SeqID 1.			
XX				
KW	unnatural base; 5'-substituted-2'-oxo(1H)-pyridin-3'-yl; photoreactive group; biotin; fluorescent molecule; 6'-substituted-2'-aminopurin-9'-yl; aptamer; gene therapy.			
XX				
XX	Unidentified.			
OS				
XX				
PN	WO2004007713-A1.			
XX				
PD	22-JAN-2004.			
XX				
PF	28-FEB-2003; 2003WO-JP002342.			
XX				
PR	17-JUL-2002; 2002JP-00208568.			
XX				
PA	(RIKE ) RIKEN KK.			
PA	(NISIC-) JAPAN SCI & TECHNOLOGY CORP.			
PA	(ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.			
P1	Hirao I, yokoyama S, Hirao M, Mitsui T;			
XX				
DR	WPI; 2004-122944/12.			
XX				
PT	Nucleosides or nucleotides comprising novel base particularly 5'-substituted-2'-oxo(1H)-pyridin-3'-yl group as base, useful in developing functional nucleic acids as reagents or therapeutics in treating diseases.			
PT				
PT	Example 4; SEQ ID NO 1; 78pp; Japanese.			
XX	This invention relates to nucleosides or nucleotides that a novel unnatural base, namely the 5'-substituted-2'-oxo(1H)-pyridin-3'-yl group as a base. Specifically, it refers to nucleic acid molecules integrated with a base substituted at the 5-position with either a photocreactive group (e.g. iodine or bromine), an alkanyl, alkynyl or amino group, biotin or a derivative thereof or a fluorescent molecule selected from fluorescein, 6-carboxyfluorescein, tetramethyl-/-carboxyrhodamine or derivatives thereof. The present invention describes preparing nucleic acids by carrying out transcription, replication or reverse transcription with a nucleic acid containing a 6-substituted-2'-aminopurin-9'-yl group as base which is applied as template for the integration of any of the nucleotides into its complementary position. Accordingly, the nucleosides and nucleotides are useful in developing functional nucleic acids including antisense DNAs and RNAs, ribozymes and aptamers as reagents or therapeutics in treating diseases by gene therapy. The produced nucleic acids are chemically stable with improved ease of amplification, replication and transcription for their preparation. This polynucleotide sequence is an RNA aptamer (RNA9A) given in an exemplification of the invention.			
PS				
XX				
CC	Sequence 98 BP; 24 A; 21 C; 24 G; 0 T; 29 U; 0 Other;			
CC				
CC	Query Match Score 80.0%; Best Local Similarity 100.0%; Pred. No. 2; Mismatches 0; Conservation 0; Matches 80; Gaps 0;			
CC				
CC	Qy 1 GGAGUGGAGGAUCAUCAGGAUAUCUGAACUCCGUUCUCCUAAACCGUUA	60		
CC	Db 1 GGAGUGGAGGAUCAUCAGGAUAUCUGAACUCCGUUCUCCUAAACCGUUA	60		
CC	Qy 61 AUUGGUUTTACGCAUAGGCCCTAAGGCACAGC 100			
CC	Db 59 AUUGGUUTTACGCAUAGGCCCTAAGGCACAGC 98			
<b>RESULT 7</b>				
AA299051	ID AA299051 standard; RNA; 80 BP.			
XX				
AC				
XX				
DT	21-JUN-2000 (first entry)			
XX				
DE	RNA aptamer #28 for binding Ras target protein.			
XX				
KW	Ras target protein; malignant tumour; signal transmission regulation; cell proliferation; cell differentiation; aptamer; inflammation; ss.			
XX				
KW				
XX				
OS	Homo sapiens.			
XX				
PN	WO2000009684-A1.			
XX				
PD	24-FEB-2000.			
XX				
PF	13-AUG-1999;			
XX				
PR	14-AUG-1998;			
XX				
PR	24-NOV-1998;			
XX				
PA	(NTSC-) JAPAN SCI & TECHNOLOGY CORP.			
XX				
PI	Yokoyama S, Hirao I, Sakamoto K;			
XX				
DR	WPI; 2000-224330/19.			
XX				
PT	Nucleic acid e.g. RNA aptamer capable of binding specifically to Ras target protein like Raf-1, useful in drug compositions to treat and diagnose malignant tumors and inflammation.			
PT				
PT	Claim 6; Page 42; 59pp; Japanese.			
XX				
CC	The invention relates to novel nucleic acids which bind specifically to the target protein of Ras, e.g. Raf-1. RNA aptamer (AA299051) based on these sequences are useful in the treatment and diagnosis of malignant tumours and inflammation. The nucleic acids can be used to formulate medicinal compositions that are useful in the treatment of malignant tumours and inflammation and for disease diagnosis by binding specifically to Ras target protein and regulating transmission of signal causing proliferation or differentiation of cells			
CC				
CC	Sequence 80 BP; 20 A; 16 C; 19 G; 0 T; 25 U; 0 Other;			
CC				
CC	Query Match Score 80.0%; Best Local Similarity 100.0%; Pred. No. 2; Mismatches 0; Conservation 0; Matches 80; Gaps 0;			
CC				
CC	Qy 1 GGAGUGGAGGAUCAUCAGGAUAUCUGAACUCCGUUCUCCUAAACCGUUA	60		
CC	Db 1 GGAGUGGAGGAUCAUCAGGAUAUCUGAACUCCGUUCUCCUAAACCGUUA	60		
CC	Qy 61 AUUGGUUTTACGCAUAGGCCCTAAGGCACAGC 100			
CC	Db 59 AUUGGUUTTACGCAUAGGCCCTAAGGCACAGC 98			

RESULT 8	RESULT 9
ADJ10066 standard; DNA; 147 BP.	AZ93076 standard; RNA; 45 BP.
XX ID AAZ93076.	XX ID AAZ93076.
AC XX	AC XX
XX AC AAZ9076;	XX AC AAZ9076;
DT 17-JUN-2004 (first entry)	DT 21-JUN-2000 (first entry)
XX Synthetic RNA OC aptamer (100-mer).	XX Ras target protein-RNA binding sequence #1 for generating primers.
DE ss; unnatural base; 5'-substituted-2'-oxo(1H)- pyridin-3-yl;	DE Ras target protein; malignant tumour; signal transmission regulation;
KW photoreactive group; biotin; fluorescent molecule;	XX cell proliferation; cell differentiation; aptamer; inflammation; ss.
KW 6-substituted-2-aminopurine-9-yl; aptamer; gene therapy.	XX OS Homo sapiens.
XX	XX
OS XX	PN WO200009684-A1.
SS XX	XX 24-FEB-2000.
PR XX	PP 13-AUG-1999; 99WO-JP004399.
PP XX	XX 14-AUG-1998; 98JP-0033284.
PR 17-JUL-2002; 2002JP-00208568.	PR 24-NOV-1998; 98JP-0033284.
XX (RIKE ) RIKEN KK.	XX (NISCC-) JAPAN SCI & TECHNOLOGY CORP.
PA (RIKE ) JAPAN SCI & TECHNOLOGY CORP.	PA (NISCC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.	XX PI Yokoyama S, Hirao I, Sakamoto K;
XX	XX DR WPI; 2000-224330/19.
DR XX	XX PT Nucleic acid e.g. RNA aptamer capable of binding specifically to Ras target protein like Raf-1, useful in drug compositions to treat and diagnose malignant tumors and inflammation.
WPI; 2004-122944/12.	XX PT Disclosure; Page 52; 59pp; Japanese.
XX	XX PS Disclosure; Page 52; 59pp; Japanese.
PS XX	XX CC The invention relates to novel nucleic acids which bind specifically to the target protein of Ras, e.g. Raf-1. RNA sequences (AA59076-299077) represent the sequence of aptamer which binds the Ras target protein which are used to generate PCR primers. Aptamers based on these sequences are useful in the treatment and diagnosis of malignant tumours and inflammation. The nucleic acids can be used to formulate medicinal compositions that are useful in the treatment of malignant tumours and inflammation and for disease diagnosis by binding specifically to Ras target protein and regulating transmission of signal causing proliferation or differentiation of cells
PS 10; 78pp; Japanese.	XX Sequence 45 BP; 11 A; 10 C; 7 G; 0 T; 17 U; 0 Other;
XX	XX Query Match 45.0%; Score 46; DB 3; Length 45;
This invention relates to nucleosides or nucleotides that a novel unnatural base, namely the 5'-substituted-2'-oxo(1H)- pyridin-3-yl group as a base. Specifically, it refers to nucleic acid molecules integrated with a base substituted at the 5-position with either a photoreactive group (e.g. iodine or bromine), an alkanyl, alkynyl or amino group, biotin or a derivative thereof or a fluorescent molecule selected from fluorescein, 6-carboxyfluorescein, tetramethyl-7-carboxyrhodamine or derivatives thereof. The present invention prepares nucleic acids by carrying out transcription, replication or reverse transcription with a nucleic acid containing a 6-substituted-2'-aminopurine-9-yl group as base which is applied as template for the integration of any of the nucleotides into its complementary position. Accordingly, the nucleosides and nucleotides are useful in developing functional nucleic acids including antisense DNAs and RNAs, ribozymes and aptamers as reagents or therapeutics in treating diseases by gene therapy. The produced nucleic acids are chemically stable with improved ease of amplification, replication and transcription for their preparation. This polynucleotide sequence is a synthetic RNA OC aptamer (100-mer) given in an exemplification of the invention.	XX Best Local Similarity 100.0%; Pctd. No. 5.3e-06;
XX	XX Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Sequence 147 BP; 32 A; 34 C; 37 G; 0 T; 44 U; 0 Other;	Qy 27 AUGUAGACUCGCGUCUCCUCUCAAACAGUAUAUAUUGAUUAG 71
Query Match 50.6%; Score 50.6; DB 12; Length 147;	Db 1 AUGUAGACUCGCGUCUCCUCUCAAACAGUAUAUAUUGAUUAG 45
Best Local Similarity 75.2%; Prod. No. 8e-08;	XX
Matches 76; Conservative 0; Mismatches 24; Indels 1; Gaps 1;	AC AAC75113;
XX	XX DT 08-FEB-2001 (first entry)
Query Match 1 GGGAGGGAGGAATCAUGAGGAUATGAGCTTCAACCAAGGUUA 59	XX Human ORF6 ORF668 polynucleotide sequence SEQ ID NO:1335.
Best Local Similarity 75.2%; Prod. No. 8e-08;	DB DE Human ORF6 ORF668 polynucleotide sequence SEQ ID NO:1335.
Matches 76; Conservative 0; Mismatches 24; Indels 1; Gaps 1;	XX KW Human; open reading frame; ORF6; detection; cyostatic; hepatotropic;
XX	XX KW pulmonary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
Query Match 1 GGGAGGGAGGAATCAUGAGGAUATGAGCTTCAACCAAGGUUA 60	DB Db 60 AAUTGGGTTAATGCAUAGCCTTGAGCAAGCTTGCTTGGUAGCTG 100
Best Local Similarity 75.2%; Prod. No. 8e-08;	DB Db 61 CACCUUAAACAGAUAAUAGCCUAGCAAGCUUG 101
Matches 76; Conservative 0; Mismatches 24; Indels 1; Gaps 1;	XX KW anticonvulsant; osteopathic; antiarthritic; immunosupresant; cardiac;
XX	XX KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;

KW hypotensive; dermatological; immunosuppressive; antilinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive; ss.

XX Homo sapiens.

XX WO20058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008621.

XX PR 31-MAR-1999; 99US-0127607P.

XX PR 05-APR-1999; 99US-0127636P.

XX PR 05-APR-1999; 99US-0127728P.

XX PR 30-MAR-2000; 2000US-00540763.

PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX DR WPI; 2000-602362/57.

XX DR P-PSDB; AAB40904.

XX PS Page 1146-1147; 5507pp; English.

PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for creating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX

CC AAC7446 to AAC77306 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORF-X open reading frames 1 to 3161. The ORF-X  
 CC sequences have activities such as: cytosstatic; hepatotropic; vulneracy;  
 CC antisporotrophic; nocotoric; neuroprotective; osteopatric;  
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 CC cardiant; thrombolytic; coagulant; vasoconstrictor; hypotensive;  
 CC dermatological; immunosuppressive; antilinflammatory; antibacterial;  
 CC antiviral; antifungal; antirheumatic; antithyroid; and antiaemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathologic conditions associated with an  
 CC ORF-X-associated disorder. The nucleic acids can be used to express ORF-X  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX Sequence 748 BP; 134 A; 218 C; 221 G; 174 T; 0 U; 1 Other;

SQ Score 30.6%; Score 30.6%;  
 Best Local Similarity 46.8%; Pred. No. 1.6;  
 Matches 36; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

Qy 18 UCGAGGCAUAGUGACAUUCGUUCUCCUAAACCAGUAUAAAUGGUUAGAUAGUAUG 77

Db 317 TTGGGTTCACCTGAGCGTGTAGTCAGACCGTTAGTCATG 258

Qy 78 CCUTAGCGCAGCGAGC 94

Db 257 CCTGGGCCGGGAGC 241

RESULT 11  
 ABN24508/C  
 ID ABN24508 standard; cDNA; 748 BP.  
 XX  
 AC ABN24508;  
 XX DT 24-JUN-2002 (first entry)  
 DE Human ORF-X polynucleotide sequence SEQ ID NO:17493.  
 KW Human; open reading frame; ORF-X; gene therapy; cancer; Cirrhosis;  
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
 KW hypertension; hypothyroidism; cholesterol ester storage disease;  
 KW immune deficiency; immune disorder; infectious disease;  
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
 KW myasthenia gravis; Gene; ss.  
 XX OS Homo sapiens.  
 XX PN WO20019253-A2.  
 XX PR 29-MAY-2000; 2001WO-US010836.  
 XX PR 30-MAY-2000; 2000US-0206132P.  
 XX PR 29-AUG-2000; 2000US-0226716P.  
 XX PD 06-DEC-2001.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PI Shimkets RA, Leach MD;  
 XX DR WPI; 2002-106308/14.  
 XX DR P-PSDB; ABP08756.  
 XX PS Disclosure; SEQ ID NO 17493; 1037pp; English.  
 XX PT Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and autoimmune disorders.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PI Shimkets RA, Leach MD;  
 XX DR WPI; 2002-106308/14.  
 XX DR P-PSDB; ABP08756.  
 XX PS Disclosure; SEQ ID NO 17493; 1037pp; English.  
 XX CC The present invention describes substantially purified human proteins  
 CC (referred to as open reading frame, ORF-X, where X is 1-1149) (see Table 1  
 CC in the specification). ABN1762 to ABN2752 encode the human ORF-X  
 CC proteins given in ABP00010 to ABP11500. ORF-X proteins are useful for  
 CC treating or preventing a pathology associated with an ORF-X-associated  
 CC disorder in humans, and in the manufacture of a medicament for treating a  
 CC syndrome associated with an ORF-X-associated disorder. ORF-X Polynucleotide  
 CC sequences can be used in gene therapy. ORF-X sequences can be used in the  
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
 CC benign tumours, keloid, degenerative disorders, hemorrhage,  
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
 CC storage disease, various immune deficiencies and disorders, infectious  
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 CC disease and autoimmune inflammatory eye disease. ORF-X proteins are also  
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 CC bone degenerative disorders, or periodontal disease, and for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC CC reperfusion injury in various tissues and conditions resulting from  
 CC systemic cytokine damage. N.B. The sequence data for this patent did not  
 CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)  
 XX SQ Sequence 748 BP; 134 A; 218 C; 221 G; 174 T; 0 U; 1 Other;  
 Query Match 30.6%; Score 30.6%;  
 Best Local Similarity 46.8%; Pred. No. 1.6;  
 Matches 36; Conservative 12; Mismatches 29; Indels 0; Gaps 0;





GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 30, 2004, 11:12:16 ; Search time 51.1111 Seconds  
 (without alignments)  
 1390.674 Million cell updates/sec

Title: US-09-529-397C-25

Perfect score: 100

Sequence: 1 gggttagggggaaatcaucg.....uaggcacacaaggcaucugc 100

Scoring table: IDENTITY NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters:

1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing First 45 summaries

Database : Issued Patents NA:\*

1: /seqn2\_6/.ptoodata/1/ina15A\_COMB.seq:\*\*  
 2: /con2\_6/.ptoodata/1/ina15B\_COMB.seq:\*\*  
 3: /seqn2\_6/.ptoodata/1/ina16A\_COMB.seq:\*\*  
 4: /seqn2\_6/.ptoodata/1/ina16B\_COMB.seq:\*\*  
 5: /seqn2\_6/.ptoodata/1/ina1PCUTUS\_Comb.seq:\*\*  
 6: /seqn2\_6/.ptoodata/1/ina1backtus\_Comb.seq:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
C 1	30.4	30.4	366	4	US-09-328-352-608	Sequence 608, App
C 2	29.8	29.8	4285	3	US-09-410-464-1	Sequence 1, Appli
C 3	27.4	27.4	405	4	US-09-270-767-6441	Sequence 6441, Ap
C 4	27.4	27.4	405	4	US-09-270-767-21723	Sequence 21723, A
C 5	27.4	27.4	405	4	US-08-675-773B-3	Sequence 3, Appli
C 6	27	27	3791	3	US-09-513-999C-21500	Sequence 21500, A
C 7	26.8	26.8	415	4	US-09-513-999C-26061	Sequence 26061, A
C 8	25.8	25.8	798	4	US-08-956-171E-162	Sequence 162, App
C 9	25.8	25.8	798	4	US-08-781-986A-162	Sequence 162, App
C 10	25.6	25.6	372	4	US-08-956-171E-798	Sequence 798, App
C 11	25.6	25.6	372	4	US-08-781-986A-798	Sequence 798, App
C 12	25	25.2	355	4	US-09-513-999C-8830	Sequence 8830, Ap
C 13	25.2	25.2	786431	4	US-09-751-3899-3	Sequence 3, Appli
C 14	25	25.0	246	4	US-09-513-999C-35958	Sequence 35958, A
C 15	25	25.0	960	4	US-09-328-352-1916	Sequence 1916, Ap
C 16	25	25.0	1830121	4	US-09-557-884-1	Sequence 1, Appli
C 17	25	25.0	1830121	4	US-09-633-990A-1	Sequence 1, Appli
C 18	25	25.0	1830121	4	US-10-329-960-1	Sequence 1, Appli
C 19	24.8	24.8	2775	4	US-09-248-796A-4695	Sequence 4695, Ap
C 20	24.8	24.8	3786	3	US-08-975-762-42	Sequence 42, Appli
C 21	24.8	24.8	3786	3	US-09-295-029-42	Sequence 42, Appli
C 22	24.8	24.8	3786	3	US-09-106-582-42	Sequence 42, Appli
C 23	24.8	24.8	3786	4	US-09-159-469-42	Sequence 42, Appli
C 24	24.8	24.8	3786	4	US-09-693-542-42	Sequence 42, Appli
C 25	24.8	24.8	7399	2	US-08-418-848A-9	Sequence 9, Appli
C 26	24.8	24.8	9709	3	US-08-188-583-5	Sequence 5, Appli
C 27	24.8	24.8	9709	3	US-08-388-353-1	Sequence 1, Appli

#### ALIGNMENTS

RESULT 1	US-09-328-352-608/c	Sequence 608, Application US/09328352
		; Patent No. 6562958
		; GENERAL INFORMATION:
		; APPLICANT: GARY L. BRETON et al.
		; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
		; FILE NUMBER: GTC99-03PA
		; CURRENT APPLICATION NUMBER: US/09/328,352
		; CURRENT FILING DATE: 1999-06-04
		; NUMBER OF SEQ ID NOS: 8252
		; SEQ ID NO: 608
		; LENGTH: 366
		; TYPE: DNA
		; ORGANISM: Acinetobacter baumannii
		; US-09-328-352-608
Query Match	30.4%	Score 30.4;
Best Local Similarity	39.8%	Pred. No. 0.088;
Matches 35;	Conservative 17;	Mismatches 36;
		Indels 0;
		Gaps 0;
Qy	1 GGAGAUGGAGGAATCAUCAGGGAUUAUGCGACUCCGCUCCUAAACCAAGUADA 60	
Db	101 GGTTGGTTATGATATAGCTTCTATATGTGCCACGAGTAATGCAAGTAATG 42	
Qy	61 AUUGGTTUUTAGCAUAGCCUTAGCTTCTATATGTGCCACGAGTAATGCAAGTAATG 88	
Db	41 TCTGGTTATGCGTATATTCTAGATTCA 14	
RESULT 2	US-09-410-464-1	
		; Sequence 1, Application US/09410464
		; Patent No. 6395892
		; GENERAL INFORMATION:
		; APPLICANT: STRAUSS et al.
		; TITLE OF INVENTION: floral homeotic genes for manipulation of flowering in poplar and other plant species.
		; FILE REFERENCE: 53375
		; CURRENT APPLICATION NUMBER: US/09/410,464
		; CURRENT FILING DATE: 1999-10-01
		; EARLIER APPLICATION NUMBER: 09/287,700
		; EARLIER FILING DATE: 1999-04-05
		; EARLIER APPLICATION NUMBER: 60/080,851
		; EARLIER FILING DATE: 1998-04-06
		; NUMBER OF SEQ ID NOS: 24
		; SOFTWARE: Patentin Ver. 2.0
		; SEQ ID NO: 1
		; LENGTH: 4285

TYPE: DNA  
 ORGANISM: Populus balsamifera subsp. trichocarpa  
 US-09-410-464-1

Query Match Score 29.8%; Pred. No. 0.33; Length 4285;  
 Best Local Similarity 33.3%; Matches 32; Indels 0; Gaps 0;  
 Matches 27; Conservative 22; Mismatches 32;

Qy 8 GAGGAUUDAUUGGGCAUAUGGACUGCGUCUCCUAAACCAUGUAUAAAUCGU 67  
 Db 1193 GATTAAATCTGTAACTTCTTCGTTATGCCCTCAATCCATCTATTGTT 1252

Qy 68 UUAGCAUAGGCCUTAGGCACA 88  
 Db 1253 TTGATGTTCTTGTAGATACA 1273

RESULT 3  
 US-09-270-767-6441  
 Sequence 6441, Application US/092270767  
 Patent No. 6703491  
 GENERAL INFORMATION:  
 APPLICANT: Homburger et al.  
 TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
 FILE REFERENCE: File Reference: 7326-03-04  
 CURRENT FILING DATE: 1999-03-17  
 NUMBER OF SEQ ID NOS: 62517  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO: 6441  
 LENGTH: 405

ORGANISM: Drosophila melanogaster  
 US-09-270-767-6441

Query Match Score 27.4%; Pred. No. 1.2; Length 405;  
 Best Local Similarity 35.5%; Matches 41; Indels 0; Gaps 0;  
 Matches 33; Conservative 19; Mismatches 41;

Qy 7 GGAGAAUUCAUAGGGCAUAUGGACUGCGUCUCCUAAACCAUGUAUAAAUCGU 66  
 Db 2227 GTGTAATCCATTGGTCAATTGGTCCCTCTCAACTCAACACCTGGATAATTCAT 286

Qy 67 UUAGCAUAGGCCUTAGGCAGACAGCAGTUCUG 99  
 Db 2877 TTAAAAAGTCGTAAACCAGGAAGATCTG 319

RESULT 4  
 US-09-270-767-21723  
 Sequence 21723, Application US/092270767  
 Patent No. 6703491  
 GENERAL INFORMATION:  
 APPLICANT: Homburger et al.  
 TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
 FILE REFERENCE: File Reference: 7326-03-04  
 CURRENT FILING DATE: 1999-03-17  
 NUMBER OF SEQ ID NOS: 62517  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO: 21723  
 LENGTH: 405

ORGANISM: Drosophila melanogaster  
 US-09-270-767-21723

Query Match Score 27.4%; Pred. No. 1.2; Length 405;  
 Best Local Similarity 35.5%; Matches 41; Indels 0; Gaps 0;  
 Matches 33; Conservative 19; Mismatches 41;

Qy 7 GGAGAAUUCAUAGGGCAUAUGGACUGCGUCUCCUAAACCAUGUAUAAAUCGU 66  
 Db 2227 GTGTAATCCATTGGTCAATTGGTCCCTCTCAACTCAACACCTGGATAATTCAT 286

RESULT 5  
 US-08-675-773B-3/C  
 Sequence 3, Application US/08675773B  
 Patent No. 6166288  
 GENERAL INFORMATION:  
 APPLICANT: DIAMOND, LISA E  
 APPLICANT: LOGAN, JOHN S  
 APPLICANT: BYRNE, GUBERARD W  
 APPLICANT: SHARMA, ADY  
 TITLE OF INVENTION: METHOD OF PRODUCING TRANSGENIC ANIMALS  
 TITLE OF INVENTION: FOR XENOTRANSPLANTATION . . .  
 NUMBER OF SEQUENCES: 42  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BRONDY AND NEIMARK, P.L.L.C.  
 STREET: 419 SEVENTH STREET, N.W., SUITE 300  
 CITY: WASHINGTON  
 STATE: D.C.  
 COUNTRY: USA

ZIP: 20004  
 COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/675,773B  
 FILING DATE: 03-JUL-1996

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: COOPER, IVER P.

REGISTRATION NUMBER: 28,005

REFERENCE/DOCKET NUMBER: DIAMOND=1A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 3791 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-675-773B-3

Query Match Score 27.4%; DB 3; Length 3791;

Best Local Similarity 41.6%; Pred. No. 2.5;

Matches 32; Conservative 14; Mismatches 31; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Ducleart, A.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

Patent No. 6783961

RESULT 6  
 US-09-513-999C-21500/C

Sequence 21500, Application US/09513999C

Patent No. 6783961

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

Patent No. 6783961

FILE REFERENCE: 59.US2.REG  
 CURRENT APPLICATION NUMBER: US/09/513,999C  
 PRIOR FILING DATE: 2000-03-24  
 PRIORITY NUMBER: US 60/122,487  
 NUMBER OF SEQ ID NOS: 36681  
 SEQ ID NO: 21500  
 LENGTH: 208  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE: misc\_feature  
 LOCATION: 13  
 OTHER INFORMATION: v=a or c or g  
 US-09-513-999C-21500

Query Match 27.0% Score 27; DB 4; Length 208;  
 Best Local Similarity 39.8%; Pred. No. 1,3; Mismatches 35; Indels 0; Gaps 0;  
 Matches 33; Conservative 15; MisMatches 35; Indels 0; Gaps 0;

Qy 13 AUCAUCAGGCAAUAGUGACUCCGUATGGGUUUAGC 72  
 Db 203 ATTACATACCACACTGTAAGTCGTCTTAAACCCGCTTGCTCATC 144

Qy 73 AU AUGCCCTAAGGCCAACGAAAGCU 95  
 Db 143 ATATTAGACCCCCCAGCTAGCT 121

RESULT 7  
 US-09-513-999C-26061/C  
 Sequence 26061; Application US/095139999C  
 Patent No. 6783961  
 GENERAL INFORMATION:  
 APPLICANT: Dumas, Moline Edwards, J.B.  
 DUCLER, A.  
 APPLICANT: Giordano, J.Y.  
 TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 Patent No. 6783961

FILE REFERENCE: 59.US2.REG  
 CURRENT APPLICATION NUMBER: US/09/513,999C  
 CURRENT FILING DATE: 2000-03-24  
 PRIOR APPLICATION NUMBER: US 60/122,487  
 PRIOR FILING DATE: 1999-02-26  
 NUMBER OF SEQ ID NOS: 36681  
 SOFTWARE: Patent .pm  
 SEQ ID NC: 26061  
 LENGTH: 415  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE: misc\_feature  
 NAME/KEY: misc\_feature  
 LOCATION: 340  
 OTHER INFORMATION: y=c or t  
 FEATURE: misc\_feature  
 NAME/KEY: misc\_feature  
 LOCATION: 344  
 OTHER INFORMATION: k=g or t  
 US-09-513-999C-26061

Query Match 26.8% Score 26.8; DB 4; Length 415;  
 Best Local Similarity 48.4%; Pred. No. 2; Mismatches 22; Indels 0; Gaps 0;  
 Matches 30; Conservative 10; MisMatches 22; Indels 0; Gaps 0;

Qy 28 UGUGACUCGCGUUCUCCUAACCAAGUAUAAAUGGUUUTAGCAUAGCGCUUAGCGAC 87  
 Db 269 TGTAGACACCCCTGGCATACCCAGTAACCTGGCAACTGGTGCATATCCCAGAGAC 210

Qy 88 AG 89  
 Db 209 AG 208

RESULT 8  
 US-08-956-171E-162  
 Sequence 162; Application US/08956171E  
 Patent No. 6593114  
 GENERAL INFORMATION:  
 APPLICANT: Charles Kunsch  
 Michael R. Fannon  
 Gil H. Choi  
 Patrick S. Dillon  
 Craig A. Rosen  
 Steven C. Barash  
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 5256  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/956,171E  
 FILING DATE: 20-Oct-1997  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/009,861  
 FILING DATE: January 5, 1996  
 APPLICATION NUMBER: 08/781,986  
 FILING DATE: January 3, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mark J. Hyman  
 REFERENCE DOCKET NUMBER: 46,789  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (240) 314-1224  
 TELEFAX: (301) 309-8439  
 INFORMATION FOR SEQ ID NO: 162:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 788 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 162:  
 US-08-956-171E-162

Query Match 25.8% Score 25.8; DB 4; Length 798;  
 Best Local Similarity 31.2%; Pred. No. 5,7;  
 Matches 24; Conservative 21; MisMatches 32; Indels 0; Gaps 0;

Qy 22 GGCAUAGUGAGACUGCGUUCUCCUAACCAAGUAUAAAUGGUUUTAGCAUAGCGCUU 81  
 Db 196 GGCATTTTACTTCTTCTTATTTGTCAGTATTAAAGTGTATTAGGAATCTCAT 255

Qy 82 AGCGACAGCAAGCUCU 98  
 Db 256 CGCTTCTGCGATTCT 272

RESULT 9  
 US-08-781-986A-162  
 Sequence 162; Application US/08781986A  
 Patent No. 6737245  
 GENERAL INFORMATION:  
 APPLICANT: Charles Kunsch  
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 5255  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS / version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/781,986A  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Benson, Bob  
 REGISTRATION NUMBER: 30 446  
 REFERENCE/DOCKET NUMBER: PB248PP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8503  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 162:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 798 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 US-08-781-986A-162

Query Match 25.8%; Score 25.8; DB 4; Length 798;  
 Best Local Similarity 31.2%; Pred. No. 5.7;  
 Matches 24; Conservative 21; Mismatches 32; Indels 0; Gaps 0;

Qy 22 GGCAUAGUGCGACUCGGGUUCUCAAAACCAGUATAATGGGTTAGCAUAUGCCTU 81  
 Db 196 GGATCTTTCACATCTTCTATTTCTGCTAGTATTAAGCTGTATTGATTCTCAT 255

Qy 82 AGGCACAGAACGUUCU 98  
 Db 256 CGCTCTGGGATTTCT 272

RESULT 10  
 US-08-956-171E-798  
 Sequence 798, Application US/08956171E  
 Patent No. 6593114

GENERAL INFORMATION:  
 APPLICANT: Charles Kunsch  
 Gil H. Choi  
 Patrick S. Dillon  
 Craig A. Rosen  
 Steven C. Barash  
 Michael R. Fannon  
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 5256  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/781,986A  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Benson, Bob  
 REGISTRATION NUMBER: 30 446  
 REFERENCE/DOCKET NUMBER: PB248PP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 798:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 372 base pairs

Query Match 25.6%; Score 25.6; DB 4; Length 372;  
 Best Local Similarity 36.9%; Pred. No. 5.3;  
 Matches 24; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

Qy 27 AUGUGACUCCGUCCUCCUCAAAACCAGUATAAUCCGUUUTAGCAUAUGCCUAGCGA 86  
 Db 223 ATTAAACACTTCTCATCTTTCATCTTACAGTTTGATGATTGTTTCAGTTGTTGCA 282

Qy 87 CAGCA 91  
 Db 283 GGCCA 287

RESULT 11  
 US-08-781-986A-798  
 Sequence 798, Application US/08781986A  
 Patent No. 6737248

GENERAL INFORMATION:  
 APPLICANT: Charles Kunsch  
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 5255  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/781,986A  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Benson, Bob  
 REGISTRATION NUMBER: 30 446  
 REFERENCE/DOCKET NUMBER: PB248PP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 798:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 372 base pairs

Query Match Score 25.6%; DB 4; Length 372;  
 Best Local Similarity 36.9%; Pred. No. 5.3%;  
 Matches 24; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

Qy 27 AUGUCACUCCGCUCCUCUCAAACAGUAUAAAUGGUUUTAGCAUAGCGA 86  
 Db 223 ATTCAACACCCCTTTCATCCCTTTAACGGTTGCTGTGCATTGNTTTGCA 282

Qy 87 CAGCA 91  
 Db 283 GGCCA 287

RESULT 12 US-09-513-999C-8830  
 Sequence 8830, Application US/09513999C  
 Patent No. 6783961  
 GENERAL INFORMATION:  
 APPLICANT: Dumas Milne Edwards, J.B.  
 APPLICANT: Duclert, A.  
 APPLICANT: Giordano, J.Y.  
 TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 Patent No. 6783961  
 FILE REFERENCE: 59 US2 REG  
 CURRENT APPLICATION NUMBER: US/09/513,999C  
 CURRENT FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/122,487  
 PRIOR FILING DATE: 1999-02-26  
 NUMBER OF SEQ ID NOS: 36681  
 SOFTWARE: Patent.pml  
 SEQ ID NO: 8830  
 LENGTH: 355  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE: misc\_feature  
 NAME/KEY: misc\_feature  
 LOCATION: 299  
 OTHER INFORMATION: k=g or t  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 319  
 OTHER INFORMATION: r=a or g  
 US-09-513-999C-8830

Query Match Score 25.6%; DB 4; Length 372;  
 Best Local Similarity 36.9%; Pred. No. 5.3%;  
 Matches 24; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

Qy 27 AUGUCACUCCGCUCCUCUCAAACAGUAUAAAUGGUUUTAGCAUAGCGA 86  
 Db 166 ATTCAATCTCTTAAACGGTTTCTTAAACGGTTTCTTAACTGCTATAGATGTCGA 225

Qy 87 CA 88  
 Db 226 CA 227

RESULT 13 US-09-781-986A-798  
 Sequence 3, Application US/09751389  
 Patent No. 663034  
 GENERAL INFORMATION:  
 APPLICANT: GUELER, Karl et al.  
 TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

Query Match Score 25.2%; DB 4; Length 786431;  
 Best Local Similarity 42.6%; Pred. No. 91;  
 Matches 23; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

Qy 44 CUUCAAACGUUAAAUTGGUUUTAGCAUAGCGACAGCAAGCUUC 97  
 Db 507097 CTCCAGCTGTGATAATGTTTAAGCCATTGCTTAGTGACATTGTTTC 507044

US-09-513-999C-35958  
 Sequence 35958, Application US/09513999C  
 Patent No. 6783961  
 GENERAL INFORMATION:  
 APPLICANT: Dumas Milne Edwards, J.B.  
 APPLICANT: Duclert, A.  
 APPLICANT: Giordano, J.Y.  
 TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 Patent No. 6783961  
 FILE REFERENCE: 59 US2 REG  
 CURRENT APPLICATION NUMBER: US/09/513,999C  
 CURRENT FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/122,487  
 PRIOR FILING DATE: 1999-02-26  
 NUMBER OF SEQ ID NOS: 36681  
 SEQ ID NO: 35958  
 LENGTH: 246  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 137  
 OTHER INFORMATION: m=a or c  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 153  
 OTHER INFORMATION: w=a or t  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 143  
 OTHER INFORMATION: r=a or g  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 170  
 OTHER INFORMATION: b=c or g or t  
 US-09-513-999C-35958

Query Match Score 25.0%; DB 4; Length 246;  
 Best Local Similarity 43.9%; Pred. No. 7.7%;  
 Matches 18; Conservative 13; Mismatches 10; Indels 0; Gaps 0;

Qy 30 UCGACUCCGCUUCGUUACAGUAAAUGGUUUTAGCAUAGCGA 70  
 Db 182 TAGTCCTGGGCACTCTCATGCCAGTGATAATTGTTTA 222

RESULT 15  
US-09-328-352-1916/c  
Sequence 1916, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTCA9-03FA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 1916  
LENGTH: 960  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
us-09-328-352-1916

Query Match Score 25; DB 4; Length 960;  
Best Local Similarity 36.9%; Pred. No. 12;  
Matches 24; Conservative 16; Mismatches 25; Indels 0; Gaps 0;  
QY 12 AAUCAUCAGGCAUAUGUCGACUCGGTUCGUAAACCGTUAAAUGGUUAG 71  
Db 534 AATGGCTTAGCAATACGGGACTATACTGTTCCACCATATGATAAAGATATTGGC 475  
QY 72 CAUAU 76  
Db 474 CAAAT 470

Search completed: November 30, 2004, 11:53:59  
Job time : 58.1111 secs



RESULT 2  
 Db 1253 TTATGATTTCCTAGATICA 1273  
 ;  
 ; Sequence 40531, Application US/10282122A  
 ;  
 ; GENERAL INFORMATION:  
 ;  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trwick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forseyth, R.  
 ;  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: LITRA.04A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ;  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ;  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ;  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ;  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ;  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ;  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ;  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ;  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ;  
 ; PRIOR APPLICATION NUMBER: 60/167,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ;  
 ; PRIOR APPLICATION NUMBER: 60/169,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 7864  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO: 40631  
 ; LENGTH: 675  
 ;  
 ; TYPE: DNA  
 ; ORGANISM: Ureaplasma urealyticum  
 ;  
 ; US-10-282-122A-40631  
 ;  
 ; Query Match Score 28.8%; Pred. No. 11; Length 675;  
 ; Best Local Similarity 42.2%; Mismatches 22; Indels 0; Gaps 0;  
 ; Matches 27; Conservative 15; Number of SEQ ID NOS: 563  
 Qy 4 AGUGGAGGAUUCUAGGAGCAUAGUGGGACUCGGUCCUTCAACCAGTUAUAAAATU 63  
 Db 563 AGTGGAGCAATTGGCAGTCATCATCGATGTCATCGTTAAATCGTTAATATATT 504  
 ;  
 ; RESULT 3  
 ; US-09-938-842A-2908/C  
 ; Sequence 2908, Application US/0938842A  
 ;  
 ; GENERAL INFORMATION:  
 ;  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; FILE REFERENCE: SCRIP1300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938, 842A  
 ;  
 ; CURRENT FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ;  
 ; PRIOR FILING DATE: 2000-08-24  
 ;  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ;  
 ; PRIOR FILING DATE: 2001-01-16  
 ;  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ;  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO: 2908  
 ; LENGTH: 2000;  
 ;  
 ; Query Match Score 28.6%; Pred. No. 19; Length 2000;  
 ; Best Local Similarity 35.8%; Mismatches 24; Indels 0; Gaps 0;  
 ; Matches 24; Conservative 19; Number of SEQ ID NOS: 5379  
 Qy 4 AGUGGAGGAUUCUAGGAGCAUAGUGGGACUCGGUCCUTCAACCAGTUAUAAAATU 63  
 Db 759 AGTACAGATTAGCTACATATGATCGCCGTAAATTCCTGAAACTATATAATT 700  
 ;  
 ; Query Match Score 28.6%; Pred. No. 19; Length 2000;  
 ; Best Local Similarity 35.8%; Mismatches 24; Indels 0; Gaps 0;  
 ; Matches 24; Conservative 19; Number of SEQ ID NOS: 5379  
 Qy 64 GGUUUUA 70  
 Db 699 GTTTTA 693  
 ;  
 ; RESULT 3  
 ; US-09-938-842A-2908/C  
 ; Sequence 2908, Application US/0938842A  
 ;  
 ; GENERAL INFORMATION:  
 ;  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel

RESULT 5  
US-10-425-114-30801/c  
; Sequence 30801, Application US/10425114  
; Publication No. US20040034888A1  
GENERAL INFORMATION  
; APPLICANT: Liu, Jiridong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants  
FILE REFERENCE: 38-21(53313) B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 30801  
LENGTH: 586  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: UC-ZMFLB73080F09\_FLI  
US-10-425-114-30801  
Query Match 28.2%; Score 28.2; DB 16; Length 586;  
Best Local Similarity 37.1%; Pred. No. 17;  
Matches 33; Conservative 18; Mismatches 38; Indels 0; Gaps 0;  
Qy 8 GAGGAUTCAUCGAGCAUAUGUGACGUCCUCCUCAAAACCAGUUAUAAUTGGUU 67  
Db 92 GGGCATTTATGTGTATTAGGATTCCTCTGCCTCCACGCCATATAACTGGTT 33  
Qy 68 UUAGCAUUAUGCCUATAGGCACAGGAAAGCU 96  
Db 32 CCAGGAAGGCCATTAAGGAAAGGAGTT 1245

RESULT 6  
US-10-767-701-13576/c  
; Sequence 13576, Application US/10767701  
; Publication No. US2004012684A1  
GENERAL INFORMATION  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants  
FILE REFERENCE: 38-21(53315) B  
CURRENT APPLICATION NUMBER: US/10/767,701  
CURRENT FILING DATE: 2004-01-29  
NUMBER OF SEQ ID NOS: 63128  
SEQ ID NO 13576  
LENGTH: 1612  
TYPE: DNA  
ORGANISM: Sorghum bicolor  
FEATURE:  
OTHER INFORMATION: Clone ID: SCRB1-28MAY03-CLUS1303\_1  
US-10-767-701-13576  
Query Match 28.2%; Score 28.2; DB 17; Length 1612;  
Best Local Similarity 38.2%; Pred. No. 24;  
Matches 34; Conservative 17; Mismatches 38; Indels 0; Gaps 0;  
Qy 8 GAGGAUTCAUCGAGCAUAUGUGACGUCCUCCUCAAAACCAGUUAUAAUTGGUU 67  
Db 1152 GCGCATTTGTGTTGATTAGGATTCCTCTGCCTCACTGGCCGTCAC 1093

RESULT 7  
US-10-425-115-121623/c  
; Sequence 121623, Application US/10425115  
; Publication No. US20040214272A1  
GENERAL INFORMATION  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants  
FILE REFERENCE: 38-21(53222) B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 363326  
SEQ ID NO 121623  
LENGTH: 1851  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_42398C.1  
US-10-425-115-121623  
Query Match 28.2%; Score 28.2; DB 18; Length 1851;  
Best Local Similarity 37.1%; Pred. No. 26;  
Matches 33; Conservative 18; Mismatches 38; Indels 0; Gaps 0;  
Qy 8 GAGGAUTCAUCGAGCAUAUGUGACGUCCUCCUCAAAACCAGUUAUAAUTGGUU 67  
Db 1333 GGGGATTTTATGTGTATTAGGATTCCTCTGCCTCACCCCATATAACTCGTT 1274  
Qy 68 UUAGCAUUAUGCCUATAGGCACAGGAAAGCU 96  
Db 1273 CCAGGAAGGCCATTAAGGAAAGGAGTT 1245

RESULT 8  
US-10-067-514-1/c  
; Sequence 1, Application US/10067514  
; Publication No. US2003005431A1  
GENERAL INFORMATION  
; APPLICANT: Gretsasdottir, Solveig  
; APPLICANT: Jonasdottir, Sif  
; APPLICANT: Revinsdottir, Sigridur Th.  
; APPLICANT: Reyrinsdottir, Sigridur Th.  
TITLE OF INVENTION: HUMAN STROKE GENE  
FILE REFERENCE: 2345-2010-003  
CURRENT APPLICATION NUMBER: US/10/067,514  
CURRENT FILING DATE: 2003-02-04  
PRIOR APPLICATION NUMBER: US 09/811,352  
PRIORITY FILING DATE: 2001-03-19  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSEQ For Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 1691139  
TYPE: DNA  
ORGANISM: Human  
US-10-067-514-1  
Query Match 28.0%; Score 28; DB 14; Length 1691139;  
Best Local Similarity 41.7%; Pred. No. 3.6e-02;  
Matches 25; Conservative 15; Mismatches 20; Indels 0; Gaps 0;  
Qy 12 AAUUCAUUCGAGCAUAUGUGACGUCCUCCUCAAAACCAGUUAUAAUTGGUU 71  
Db 127542 AGTGTGTCATGCTTGGGGAGCTTATCCCTCAAGCAGTATAAAAGTGTTCAG 127483

RESULT 9  
US-10-419-723-1/c  
; Sequence 1, Application US/10419723  
; Publication No. US20040014099A1  
GENERAL INFORMATION  
; APPLICANT: Gretsasdottir, Solveig

1 APPLICANT: Jonsdottir, Sif ; GENERAL INFORMATION:  
 1 APPLICANT: Thorleifsson, Gudmar ; APPLICANT: David W. Morris  
 1 APPLICANT: Gulcher, Jeffrey ; APPLICANT: Marc S. Malandro  
 1 TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR HUMAN STROKE ; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
 1 TITLE OF INVENTION: METHODS OF TREATMENT ; FILE REFERENCE: 59452001000  
 1 FILE REFERENCE: 2345\_2010-005 ; CURRENT APPLICATION NUMBER: US/10/419,723  
 1 CURRENT APPLICATION NUMBER: 2003-04-18 ; CURRENT FILING DATE: 2002-12-17  
 1 PRIORITY FILING DATE: 2002-09-25 ; NUMBER OF SEQ ID NOS: 866  
 1 PRIORITY FILING DATE: 2002-09-25 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 1 PRIORITY FILING DATE: 2002-09-25 ; SEQ ID NO: 744  
 1 PRIORITY FILING DATE: 2002-09-25 ; LENGTH: 93291  
 1 PRIORITY FILING DATE: 2002-09-25 ; TYPE: DNA  
 1 PRIORITY FILING DATE: 2002-09-25 ; ORGANISM: Homo sapiens  
 1 PRIORITY FILING DATE: 2002-09-25 ; SEQ ID NO: 1 ; US-10-322-281-744

Query Match 27.8%; Score 27.8; DB 17; Length 99291;  
 Best Local Similarity 31.0%; Pred. No. 1.1e+02;  
 Matches 22; Conservative 22; Mismatches 27; Indels 0; Gaps 0;

QY 12 AAUCAUCAGGCAUAUGCGACUCCGUUUCUCAAACCAAGUUAUAGGUUAG 71  
 Db 23527 ATATATTGAGACCTGTCAAGTACTTTTGTTGCRACAAATAAACGGTTAC 23468

Query Match 28.0%; Score 28; DB 16; Length 1691139;  
 Best Local Similarity 41.7%; Pred. No. 3.6e+02; Length 1691139;  
 Matches 25; Conservative 15; Mismatches 20; Indels 0; Gaps 0;  
 Db 23467 AATTTGGTTA 23457

RESULT 12  
 US-10-097-111-47 ; Sequence 47, Application US/10097111  
 ; Publication No. US20030138771A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PELLETIER, JERRY  
 ; APPLICANT: GROS, PHILIPPE  
 ; APPLICANT: DUBOW, MICHAEL  
 ; TITLE OF INVENTION: DNA SEQUENCES FROM S. PNEUMONIAE BACTERIOPHAGE EP1 THAT  
 ; ENCODE ANTI-MICROBIAL POLYPEPTIDES  
 ; FILE REFERENCE: 073406-0603  
 ; CURRENT APPLICATION NUMBER: US/10/097111  
 ; PRIORITY FILING DATE: 2002-07-24  
 ; PRIOR APPLICATION NUMBER: 09/676,412  
 ; PRIORITY FILING DATE: 2000-09-29  
 ; PRIOR APPLICATION NUMBER: 60/157,218  
 ; NUMBER OF SEQ ID NOS: 552  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO: 47  
 ; LENGTH: 522  
 ; TYPE: DNA  
 ; ORGANISM: Streptococcus pneumoniae  
 US-10-097-111-47

Query Match 27.0%; Score 27; DB 15; Length 522;  
 Best Local Similarity 40.0%; Pred. No. 42; Mismatches 30; Indels 0; Gaps 0;

QY 13 AUCAUCAGGCAUAUGCGACUCCGUUUCUCAAACCAAGUUAUAGGUUAGC 72  
 Db 186 ATTATATTGAGACCTGTCAAGGATGACCAATCGCTTCATCAAGGATGAC 245

Query Match 27.8%; Score 27.8; DB 15; Length 1093;  
 Best Local Similarity 38.0%; Pred. No. 29; Mismatches 27; Indels 0; Gaps 0;

QY 4 AGUGGAGGAATCAUCGGCAUAUGCGACUCCGUUUCUCAAACCAAGUUAUAT 63  
 Db 213 AGTGGACTTTCTAGACAATGAAGGATCGTTAACCTTAAACGAGTTAAATC 272

QY 64 GUUUUAGCAU 74  
 Db 273 GTTGTTCAT 283

RESULT 13  
 US-10-097-111-10 ; Sequence 10, Application US/10097111  
 ; Publication No. US20030138771A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PELLETIER, JERRY  
 ; APPLICANT: GROS, PHILIPPE

RESULT 11  
 US-10-322-281-744/C ; Sequence 744, Application US/10322281  
 ; Publication No. US20040126762A1

APPLICANT: DUBOW, MICHAEL  
 TITLE OF INVENTION: DNA SEQUENCES FROM S. PNEUMONIAE BACTERIOPHAGE EP1 THAT  
 FILE REFERENCE: 0734-06-053  
 CURRENT APPLICATION NUMBER: US/10/097,111  
 PRIOR APPLICATION NUMBER: 09/676,412  
 PRIOR FILING DATE: 2002-07-24  
 NUMBER OF SEQ ID NOS: 552  
 SOFTWARE: Patentin Ver. 2.1  
 SEQ ID NO: 10  
 LENGTH: 56506  
 TYPE: DNA  
 ORGANISM: Streptococcus pneumoniae  
 US-10-097-111-Str

Query Match Score 27; DB 15; Length 56506;  
 Best Local Similarity 40.0%; Pred. No. 2. 4e+02;  
 Matches 30; Conservative 15; Mismatches 30; Indels 0; Gaps 0;  
 Qy 13 AUCAUCAGGCCAUAUUGGACTCCGUUCCGUUAAACCAGUUAUAUUGUUTAGC 72  
 Db 1535 ATTATTGACAGACTGGACACCGCTGTCTTCAGGGATGAACCAAATGCCTTGC 1594  
 Qy 73 AUAUGCCCTAUGCGAC 87  
 Db 1595 AAATGCGTTGACAC 1609

RESULT 14  
 US-10-271-181B-6  
 ; Sequence 6, Application US/10271181B  
 ; Publication No. US20030155672A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kalyanaraman Ramnarayan  
 ; APPLICANT: Edward T. Maggio  
 ; APPLICANT: P. Patrick Hess  
 ; TITLE OF INVENTION: Use of Computationally Derived Protein  
 ; TITLE OF INVENTION: Structures of Genetic Polymorphisms in Pharmacogenomics for  
 ; TITLE OF INVENTION: Drug Design and Clinical Applications  
 ; FILE REFERENCE: 24/37-1506D  
 ; CURRENT APPLICATION NUMBER: US/10/271,181B  
 ; PRIOR APPLICATION NUMBER: 09/438,566  
 ; PRIOR FILING DATE: 1999-11-10  
 ; PRIOR APPLICATION NUMBER: 09/704,362  
 ; PRIOR APPLICATION NUMBER: 09/709,905  
 ; PRIOR FILING DATE: 2000-11-10  
 ; NUMBER OF SEQ ID NOS: 118  
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0  
 ; SEQ ID NO: 6  
 ; LENGTH: 1116  
 ; TYPE: DNA  
 ; ORGANISM: Human Immunodeficiency Virus (HIV)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1) .. (297)  
 ; OTHER INFORMATION: HIV Protease

for

Search completed: November 30, 2004, 12:01:02

Job time : 237.667 secs

Db 800 TAGAGCCCTTCAGAAACAAAATCCAGACATGGCATCTCAATACATGGATGTTGT 859  
 Qy 66 UUUUAGCAUAGCCUAGCGACAGCA 91  
 Db 860 ATGAGGATCTGACTAGATAAGGA 885

RESULT 15  
 US-10-36-493-35948  
 ; Sequence 35948, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinckle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; PLANTS WITH IMPROVED PROPERTIES  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10 (52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2000-02-28  
 ; PRIOR APPLICATION NUMBER: US/10/369,493  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO. 35948  
 ; LENGTH: 1884  
 ; TYPE: DNA  
 ; ORGANISM: Mesorhizobium loti  
 US-10-369-493-35948

Query Match Score 26.6%; Score 26.6%;  
 Best Local Similarity 49.0%; Pred. No. 94;  
 Matches 24; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

Qy 26 UAUUGGACTCCGUUCCGUUAAACCAGUUAUAUUGUUTAGCAU 74  
 Db 364 TATGGGAGGCCGTGGCATCAACGGCTATGAATGCGTCGCT 412

Query Match Score 26.8%; Score 26.8%;  
 Best Local Similarity 38.4%; Prod. No. 65;  
 Matches 33; Conservative 16; Mismatches 37; Indels 0; Gaps 0;  
 ; OTHER INFORMATION: Portion of HIV Reverse Transcriptase  
 US-10-271-181B-6

Qy 6 UGGAGGAUUCGAGGCAUAUGUCACUCGGCUUCUCAAACAGUUAUAUUGG 65

